

The great Berlin greenhouse compromise

The Berlin climate conference decided correctly not to settle for a tough protocol on greenhouse emissions here and now, but negotiating a satisfactory first step will take all the time there is left.

THE famous Conference of Parties to the Rio Treaty on Climate Change (in Berlin earlier this month) came to the right conclusion after all. The conference resisted pressure to settle on the spot for stricter limits on emissions. How could it have done otherwise when, disagreements among the treaty members apart, there would have been no means of policing any agreement foisted on them? Wisely, the conference instead agreed that the signatories will negotiate a protocol to the treaty that will bind them to restrict their emissions of greenhouse gases from some date in the future. On the model of the Vienna Treaty on the Protection of the Ozone Layer, where protocols signed at Montreal and then London have gone a long way to stabilize the release of chlorofluorocarbons (CFCs) to the atmosphere, the goal must now be that the new protocol of the climate change treaty will be the first of a succession, each tighter than its predecessor. What are the chances of that happening?

Much will depend on the willingness and even the capacity of governments to take the issue seriously. The most obvious danger is that negotiations will hang fire while the distractions of everyday business turn governments' attention to other things. On past form in other contexts, they will leave too little time for the difficult issues of principle underlying the global warming business to be ironed out. The result may well be a protocol so replete with logical inconsistencies and inequities that either the treaty members will not sign it or, perhaps worse, that they will, providing a flawed foundation for the succession of protocols that must follow. The two years ahead, the time provisionally allowed for the negotiations, are therefore critical.

Data

The most basic need is for better monitoring of greenhouse gas emissions, not simply of those of carbon dioxide. As things stand, the responsibility falls on national governments to provide the organs administering the treaty with data about emissions from their own territories. National sovereignty requires no less and, as things stand, allows no more. But how accurate are these data? Nobody can know. Figures for coal extraction are readily available in most coal-mining countries, and elsewhere can be had from customs data, but the carbon content of commercial coal can vary from one source to another by at least a factor of two. Using coal production as a proxy for carbon dioxide emission, without further information, could be in error by 20 per cent. Petroleum is a bigger bugbear; who knows how

much associated methane is burned off at well-heads and refineries, or (more damagingly) is allowed to leak off unburned, constituting a more effective greenhouse gas in its own right? By the same test, the world's paddy-fields, farm animals and landfills are only shadowy sources of the same gas, while the world's stock of motor vehicles is an even more uncertain source of nitrous oxide.

Monitoring

In present circumstances, there is no reason to suppose that governments are less than well-intentioned in their declarations (although their dilatoriness is hardly a good sign). But when the unavoidable costs of reducing greenhouse emissions increase, as is certain, there will be powerful incentives to cheat. Then, cheating (or the suspicion that others are) will be rife, and support for whatever protocol is negotiated will be undermined by recriminations. That is the strongest argument for the creation of an effective monitoring regime before restriction on emissions begin to bite. Logic requires that it should cover all the known greenhouse gases. One of the important goals is to determine not only the total emissions affecting climate forcing, but the uncertainties that inevitably attend them. Of necessity, the long-term purpose of such a system requires that there should be provision for external verification of nationally declared emissions, on the pattern of many arms-control treaties. A little swallowing of pride on sovereignty will be needed even at this early stage.

The other urgent issue is the basis on which national emissions are compared with each other. The Rio treaty lists developed countries in 'Annex I' and others in 'Annex II'. The Berlin meeting amply demonstrated that these distinctions are artificial and potentially divisive. Better to abolish them, and instead work out a basis on which the burdens of heading off threatened global warming can be equitably shared. That is not to say that countries at an early stage of their development should be subject to rigorous restrictions at this stage, but is simply to acknowledge that now-poor countries may be rich one day. Even as things are, it is a needless assault on the atmosphere that Annex II states should needlessly pollute the atmosphere for lack of the capital to use fuel efficiently.

The European Union (EU), which has its heart in the right place on global warming, could perform a public service on this potentially divisive issue. The conflict of interest between the very rich and the very poor countries of the world is mirrored within the EU by the disparate industrial



development of members such as Germany at one extreme and Portugal and Ireland at the other. Not only are these countries members of the Rio treaty in their own right, but they are also covered by the membership of the EU as a whole. But Spain has been especially resentful of the idea that the same percentage reductions of emissions should apply to all EU members. The result is that, even if several individual members states meet the obligation to emit no more greenhouse gases in 2000 than they did in 1990, the EU's overall target may be breached. The solution would be for the EU to take the lead in working out the principles of an equitable deal among its own members. The prize would be a demonstration to the rest of the world that the job can indeed be done. But even a demonstration that the task is not possible would be valuable. Then we should all know that we shall have to find some other instrument than the Rio treaty to keep global warming at bay. □

Patents for what genes?

HUGO cannot wring its hands on gene patents without declaring itself on the goals of research.

THE Human Genome Organization (HUGO) is properly concerned about the rewards the patent system offers to those who make discoveries in molecular genetics (see page 751), but it has not grasped the nettle it wishes to extirpate. HUGO's argument falls into two parts. First, it holds that fragments of genes should not be patentable if nothing is known of their function or even of the potential usefulness of a knowledge of that function or of their structure. Second, HUGO records an expression of its regret that the hard work of identifying the function of a tagged but otherwise unknown gene is likely, in present patent law, to be unpatentable. It would be "ironic and unfortunate", last week's statement says, if the patent system were to "reward the routine" investigators "while discouraging the innovators". The sentiments are impeccable, but the reality may be different from what HUGO implies.

On the first issue, though, there is no difficulty. The first application for patent protection for gene fragments was three years ago, when the US National Institutes of Health (NIH) applied for patent protection on behalf of a few hundred of J. Craig Venter's "expressed sequence tags" (ESTs), which are essentially copies in DNA sequence of the ends of the RNA molecules expressed as messengers in tissue cells. At the outset, nothing was said about the functions of the genes concerned, so the test of utility that patent applications must satisfy could not be sustained. In the end, the lawyers' argument that ESTs are indubitably artefacts, and that they are at least useful for fishing out of the genome the genes concerned, did not prevail with the patents examiners. The applications were refused the first time round, NIH withdrew its other pending application and there is now a general understanding that ESTs are not patentable.

The reasons why this state of affairs has arisen neverthe-

less deserve more attention than HUGO has given them. The use of an EST for fishing out the cognate gene from single-standard DNA is no different from techniques standard in molecular biology since the early 1970s. In patent examiners' language, it is "obvious". And otherwise, in the absence of further knowledge about the gene, it is useless. So there is logic underpinning the conclusion that ESTs as such do not qualify for patent protection. There is a case for asking that patent offices should be more rigorous in their application of these argument in the future. HUGO's statement should help to push them in that direction.

The second part of HUGO's argument is the one that really matters. Ever since excitement about human genome projects surfaced 15 years ago, it has been (or should have been) plain that the difficulty of constructing the sequence of the human genome would be dwarfed by that of telling what each of perhaps 100,000 genes do. Identifying the gene responsible for Huntington's disease, for example, took a decade, but success has unleashed a host of investigations into the link between mutations of the gene and the causation of the disease. Yet understanding, let alone prophylaxis or palliation, is a long way off.

HUGO's worry is that the patent system will reward those who happen to spot some stage, in a continuing investigation such as this, at which a drug or some other treatment can be developed, and that the people whose imagination and effort have guided the process will be, by comparison, neglected. But it has always been thus. It is unlikely that the inventor of the first mousetrap had a profound understanding of the biology of small mammals. The man who made a fortune by manufacturing the glass reflectors that mark out traffic lanes on British highways is unlikely to have been well versed in the theories of reflection and refraction. It is commonplace that the people who grab the patents on inventions arising from collaborative investigations are people who have had the wit to peel themselves off into separate corporations in good time or are companies far-sighted enough to have supported imaginative projects in return for a first-refusal undertaking.

So is there nothing to be done to assuage the sense of unfairness engendered by these happenings? On the face of things, very little. Some decades ago, it would have seemed only natural to people working in an academic environment that they should do the imaginative work and that others should reap the money rewards. There were even theories to rationalize this state of affairs as an effective spur to what is now called 'technology-transfer'. But times have changed. Many in the academic community have acquired, by the example of their peers, a vivid interest in becoming millionaires. External pressures on the academic community have also pushed people in that direction. Moreover, pharmaceutical companies (among others) are increasingly major supporters of some academic laboratories. If the pharmaceutical industry were the sole supporter of academic biomedical research, HUGO's second problem would melt away. The trouble, then, is that there would be no support for the academic investigations of the human genome that are, for many, its chief interest. □

HUGO and HGS clash over 'utility' of gene sequences in US patent law

London. The Human Genome Organization (HUGO), has thrown down the gauntlet to those who, it claims, are threatening to undermine the patentability of human genes either by making excessive demands for patent protection of gene sequences or by claiming that all such protection is immoral.

The first target includes groups such as Human Genome Sciences (HGS) in Rockville, Maryland, which is seeking to patent both partial and complete gene sequences, even where their respective biological activity has not yet been identified.

The second threat, according to HUGO officials, is exemplified by the recent decision of the European Parliament to reject attempts by the European Commission to formalize the conditions under which human genes can be patented.

The challenge from HUGO — the international body responsible for coordinating efforts to sequence the complete human genome — comes in a policy statement on the patenting of DNA sequences drawn up by a working party headed by Thomas Caskey, recently appointed head of basic research for the US pharmaceutical company Merck, and president of HUGO.

Behind the statement lies concern that public debates on gene patents are sometimes based on "misinformation concerning the underlying science".

Caskey, for example, says that he was "shocked" by the European Parliament's decision, particularly as it came shortly after a meeting in Paris — attended by a number of members of the parliament — at which "there was a general consensus that there is

no serious objection to the patenting of functioning genes".

The HUGO statement, which has been endorsed by the organization's full council, argues that patent protection of human genes is essential to create "necessary incentives for the ongoing development of products without interfering unduly with scientific research".

Even those council members who admit to unease over the idea of patenting human genes say that they acknowledge the need for flexibility. "Obviously one has to look for a compromise," says John Sulston, director of the UK Medical Research Council's Sanger Centre in Cambridge.

More contentious, at least within the biotechnology industry, is HUGO's firm stand against patenting nucleotide sequences obtained by straightforward sequencing techniques before their biological function has been worked out.

The statement claims that it is the latter work that contains the key intellectual contribution. "The task of identifying biological functions of a gene is by far the most important step in terms of both its difficulty and its social benefit," it says. "It therefore merits the most incentive and protection."

Applying for patent protection on

sequences alone — for example on either Expressed Sequence Tags (ESTs), used to identify cDNA clones, or on the cDNAs themselves — without knowing their biological function is "like applying for patents on the table of elements", says Caskey.

But this is sharply disputed by officials at HGS, which has a number of applications on sequences pending with the US Patent Office. Bill Haseltine, the company's chief executive officer, says that he welcomes HUGO's statement as a recognition of "the importance of gene patents to the betterment of human health".

But he remains strongly opposed to attempts to limit patents on partial or full gene sequences whose biological function remains unknown, claiming that even ESTs can have clear usage — for example, to determine whether a gene is being expressed or to locate a restriction enzyme — even if this use is not strictly biological.

At present, says Haseltine, HGS has 70 patents pending on full-length gene sequences of "proposed medical utility". He adds that "a substantial number are biologically active" — acknowledging that for some, no such activity has yet been defined.

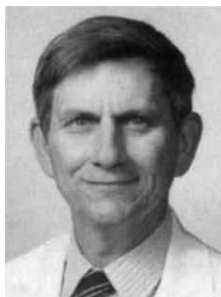
In direct conflict with Haseltine's position — and reflecting last year's decision by the US National Institutes of Health not to seek patents on cDNAs — HUGO maintains the granting of such patents could jeopardize its prime objective, namely that the financial rewards should go to those who succeed in identifying functioning genes.

The concern of HUGO officials is that patents granted on partial and uncharacterized cDNA sequences would reward those it describes as making "routine discoveries", while scuttling later patent applications by those who establish the biological function of gene sequences or the application of such knowledge. "Such an outcome would impede the development of diagnostics and therapeutics, which is clearly not in the public interest," says the policy statement.

HGS has been criticized for the conditions which, following a \$125-million deal with Merck's rival SmithKline Beecham, it has placed on access to sequences available in its databanks and those of its Institute for Genomic Research (*Nature*, 373, 376; 1995).

But Haseltine still disagrees strongly with the thrust of the HUGO statement, which he says misunderstands the patent system. "Patents are about use, and not about knowledge," he says. "I would claim that the patentability of a gene or a gene sequence is dependent on its utility, not on the advancement it makes in biological knowledge related to its function."

David Dickson



Caskey: seeks to reward creativity.

Curie laid to rest with France's heroes

Paris. Marie Curie, the co-discoverer of radioactivity, last week became the first woman to be given France's highest honour, burial in the national mausoleum, the Panthéon. The ceremony, broadcast live on national television, was attended among others by outgoing President François Mitterrand, and Lech Walesa, president of Poland, where Curie was born Marya Skłodowska in 1867.

At the ceremony, also attended by two other French Nobel prizewinners, Georges Charpak and Pierre-Gilles de Genne, Mitterrand emphasized the priority given to fundamental research under his presidency. Marie Curie won

the Nobel prize for physics in 1903 — the first woman to do so — for the discovery of natural radioactivity along with her husband Pierre, whose ashes accompanied her to the Panthéon at the request of her relatives. □

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Gamma

Study proposes PhD reforms, but no quota

Washington. The US system for educating graduate students is the best in the world and is not in need of drastic repair, despite a recent wave of complaints from young scientists, according to a report from the National Academy of Sciences.

The report, prepared by the academy's Committee on Science, Engineering and Public Policy (COSEPP), suggests new block grants to support graduate students, and more intensive career guidance. But it rejects any effort to limit the number of students embarking on PhDs, or to restrict the number of foreigners amongst them.

Phillip Griffiths, director of the Institute for Advanced Study at Princeton University, New Jersey, and chairman of the committee, says that the PhD system is not in crisis, and that its problems are more subtle than is often supposed. "We were prepared to be more radical," says Griffiths. But after listening to the people involved and studying the data, the panel "couldn't agree with the people who say the sky is falling".

Before the panel began its work, says Griffiths, "we thought PhDs were not getting jobs. They are, and they're not driving taxis." The panel found that most PhDs who failed to find academic posts eventually prospered in other sectors. "But the transition can be painful, and nobody has been

telling the students to expect it," he says.

Those who have been emphasizing the difficulties faced by PhDs in finding suitable academic positions are disappointed with the findings. "The report is little more than a collection of platitudes", says Bob Park, director of public affairs at the American Physical Society. "Everyone knows there's a problem, and all they've done is restate it, without offering any solution."

The academy recommends a national effort — including a database kept by the National Science Foundation — to provide better guidance on "realistic career options" for graduate students. It also suggests that government agencies and private foundations should provide new, block grants to fund education and training at universities and departments.

These would directly support students doing PhDs, alleviating their customary reliance on research assistantships. The panel says such direct support would help to shorten PhD courses and strengthen their educational component.

Finally, the report calls for a more flexible PhD structure to allow students who are not planning a career in academic research to take at least two alternative "pathways" to the traditional PhD. One would end with a master's degree, while the other would offer

a full PhD, combining more taught content with a shorter dissertation.

The academy rejects capping the number of PhDs awarded in the United States. There are now 25,000 each year and the number is growing. More PhDs are useful to the nation, it argues, and capping would in any case be ineffective in the short term, hard to orchestrate, and insensitive to the unpredictable variance in demand for PhDs in different fields.

Rejecting limits on foreign students, the panel argues — simultaneously and paradoxically — that the best way to strengthen the US economy, and that improving economic conditions in students' home countries will encourage most to return.

The report is, in effect, the science establishment's official response to a growing wave of anger from young scientists who find themselves trapped in low-paid, post-doctoral positions, unable to get permanent work in academic research.

Zachary Levine, recently elected to the council of the American Physical Society on this issue, says: "There are 700 jobs each year for physicists in academic research and 1,400 people being trained to fill them. We either need to cut down the number of PhDs or work out how to place them in non-academic settings."

Colin Macilwain

Now chemists hit at Smithsonian's 'anti-science' exhibit

Washington. Leaders of the American Chemical Society (ACS) will meet officials of the Smithsonian Institution in Washington within the next few days to agree on a procedure for changing the Smithsonian's *Science in American Life* exhibition to meet criticism that it has an anti-science bias.

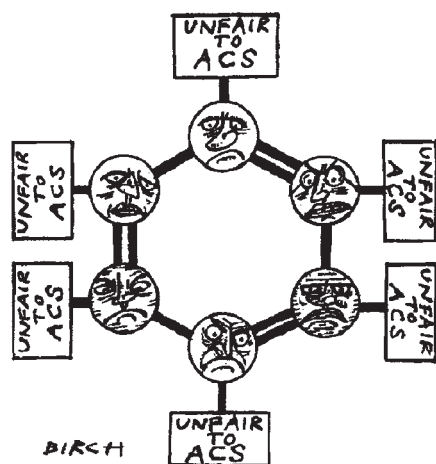
The ACS has also criticized the Smithsonian for listening to the objections of "vocal groups" who had no responsibility for the exhibition. This is a veiled reference to recent protests by the American Physical Society (APS), which has publicly accused the show of exaggerating science's failures and trivializing its accomplishments (see *Nature* 374, 207; 1995).

The meeting was arranged after the ACS — which provided \$5.53 million to sponsor the show — broke almost a year of silence and launched a vehement attack on both its content and the consultative process leading up to its design.

According to a letter sent on 21 February from Paul Walter, chairman of the society, to Michael Heyman, secretary of the Smithsonian, the four years spent planning the show were a source of "constant frustration" to ACS officials.

"Smithsonian staff consistently ignored the advice given them by an advisory board we jointly put together," wrote Walter. "We

expected the opportunity to negotiate; what we received was one rebuff after another. Arrogance and high-handedness too often



characterized the behaviour of some of the museum personnel assigned to this project."

The letter echoes criticisms made last year by the APS. It says that parts of the exhibition "demonstrate a strong built-in bias against science and display a tendency to revise and rewrite history in a 'politically correct' fashion".

Without naming the APS, the letter

implicitly attacks the rival group for its role in publicly criticizing the exhibition and proposing changes to it. "We have learned that individuals or groups who had little or nothing to do with the design and planning of the exhibit, and who made no financial contribution to it, may now expect to play some major part in its redesign," it says.

"Frankly, we worry that the Smithsonian Institution, beset as it is by public and Congressional criticism, may be tempted to make changes based on just the views of the most vocal groups presenting themselves at your doorstep."

According to Walter, the planned meeting of four senior ACS officials with four Smithsonian officials, including Heyman and the exhibition curator, Art Molella, will ensure that both sides will fully support any agreed process to modify the show. The chemical society has set up a five-member task force, chaired by Joan Shields of Long Island University, New York, to work with the Smithsonian on that process.

Ned Heindel of Lehigh University, Bethlehem, Pennsylvania, a member of the task force and a past president of ACS, promises that the chemists' group will consult the physicists during the process. "It doesn't benefit either of us to be in an arm-wrestling match over this," he says.

C.M.

US firms tipped for contract to run UK physics laboratory

London. The running of Britain's National Physical Laboratory (NPL) is likely to be taken on by a US company when the management contract for the laboratory is awarded to the private sector for the first time later this summer.

Two US companies — EDS Scicon, the software giant, and Brown and Root, the engineering contractors — are believed to be favourites to manage a laboratory that, among its many achievements, was the birthplace of radar, Barnes Wallis's Second World War 'dam-busting' bouncing bombs, and one of the world's earliest computers.

Other bidders to run an institution that has since become an international authority on measurement standards and metrology include the nuclear division of Rolls Royce, W.S. Atkins, a British civil engineering company, and Serco, a company that manages laboratories and research establishments.

Britain's Department of Trade and Industry (DTI) will select a contractor after the present shortlist of five is whittled down to two bids by May. The contract is to be announced by June or July, with the successful bidder taking over in October after an interim hand-over period.

The DTI, which is NPL's biggest customer, has promised to guarantee around £28 million-worth of work each year for the duration of the contract. Peter Clapham, the laboratory's chief executive, says he is confident of winning "substantial additional funds after competitive tendering", as well as other DTI contracts.

The proposed changes have nonetheless attracted considerable criticism. Britain's opposition Labour Party's trade and industry spokesman says he was "speechless" when he heard the news. "I cannot see the sense in this," says Lewis Mooney, MP. "There are some things you can privatize; but a national laboratory dedicated to excellence in standards and measurement is not one of them," Mooney adds.

The laboratory's 750-strong staff, most of whom are scientists, were told of NPL's impending sell-off last Christmas. Most are apprehensive about the future of jobs, intellectual property rights and the transfer of pensions in the newly privatized laboratory.

The recruitment of permanent staff has been in abeyance for more than a year, pending the outcome of the transition to private-sector management. In addition, staff numbers are due to be cut from 611 to 540 by the autumn of 1995, reflecting the drop in income from the DTI over the past three years from £40 million to less than £30 million annually.

Many staff are concerned at the principle of a prized national asset being sold over-

seas. Some also fear privatization — which is referred to officially as 'contractorization' — will result in employees currently engaged in basic research being required to switch to more lucrative short-term contract work.

Tony Mansfield, NPL's representative of the Institute of Professionals, Managers and Specialists (IPMS), the labour union that represents many researchers at the laboratory, claims that privatization could compromise NPL's world-class reputation for scientific excellence. "How will it be possible to continue long-term research when we shall all be working on short-term contracts?" he asks.

But Clapham denies that contractorization will turn the laboratory into a glorified workshop.



Clapham: says NPL will maintain its status.

"We were always vigilant of this," he says. "But ministers have declared that our status as a centre of excellence will not be impaired." Confirming that NPL would continue developing its calibration work, Clapham refuses to be drawn on whether the laboratory's new management would be obliged to continue paying for its other long-term — and potentially loss-making — research activities. "Bidders should also assume that a programme of strategic research will continue to be funded from overheads," he says.

Queen Victoria gave the NPL's original 22-acre site in Teddington near London to the Royal Society in 1900. Sir Michael Atiyah, the society's current president, says it is essential that the DTI's contract with NPL's new management should guarantee NPL's scientific excellence.

"We were not in favour of this change," says Atiyah. "But now that the government has decided to go down this road, it is important that NPL's independence and quality of research is not sacrificed."

Mansfield says staff are further irritated by the fact that there was no perceivable need for NPL's management to change hands. It was a decision "motivated more by politics than by science", he says. The proof, he says, lies in the fact that the DTI will still let NPL go, even if none of the bids is deemed satisfactory. "The present management have told us clearly that a decision on awarding the contract will shift to Michael Heseltine, [President of the Board of Trade], if none of the bids is good enough."

Ehsan Masood

Livermore hitches a ride on state lobbying efforts

San Francisco. Lawrence Livermore Laboratory, prohibited from lobbying government directly, is relying on the efforts of city and government lobbyists to win the \$1.1 billion which it is seeking to build the National Ignition Facility (NIF).

The attendance of laboratory officials at meetings on Capitol Hill has raised concerns that the laboratory is violating strict lobbying restrictions on government-funded institutions. But laboratory, company and city officials say such a presence is not only legal, but vital to ensure that members of Congress are made aware of the nuclear weapons facility that will use lasers to induce self-sustaining fusion.

"People involved in the design [of NIF] at the laboratory have the most detailed information about the facility," says Kathryn Thorpe, director of the Washington office of San Diego-based General Atomics Corp., which employs a consultant to promote the project. She says that the consultant, Pat Fulton, often invites laboratory officials to provide technical information in meetings with congressional staff and interest groups.

The Livermore laboratory would not be able to arrange such meetings on its own, as it is prohibited from promoting its projects to legislators. But officials are allowed to give technical briefings, when invited.

Laboratory representatives stress that they have been careful to avoid any appearance of lobbying. Mike Campbell, associate director for lasers at Livermore, said he is often called on to respond to requests for information from both sides of the debate about the last facility.

Laboratory representatives have held meetings both with labour unions who are anticipating more jobs from the facility, and environmental and anti-nuclear groups, who are opposed to it as a nuclear project. He said the discussions had provided an opportunity to air concerns on both sides.

Campbell admits that he works closely with Fulton, the General Atomics' lobbyist, who sometimes arranges meetings for him. He says the laboratory is delighted that General Atomics has taken on a strong advocacy role against what he describes as a well-organized campaign against the facility, and has encouraged other supporters to write to their congressmen as well.

The City of Livermore also has been active in lobbying for the giant laser. Cathie Brown, the city's mayor, says the community is keen to preserve high-paying professional jobs that have provided important side benefits such as strong science education in schools. Because the laboratory can't lobby, "the city jumped in front of the whole lobbying effort," Brown said.

Sally Lehman

Critics attack hasty review of quake prediction efforts

Tokyo. A rushed internal review of Japan's earthquake prediction programme in the light of the Kobe disaster has stimulated several leading Japanese scientists to call for a new external assessment of the country's earthquake prediction programme.

The review was carried out by the earthquake prediction subcommittee of the Geodetic Council — an advisory body to the Ministry of Education, Science and Culture. The council approved and submitted the subcommittee's recommendations to the ministry last week.

But Hideki Shimamura of Hokkaido University's Laboratory for Ocean Bottom Seismology, who is a member of Japan's Geodetic Council's earthquake prediction subcommittee, complains that the review "lacks modesty". He says that it makes no attempt to explain the present status and limitations of earthquake prediction to the general public.

Furthermore, he says it was agreed by researchers and government officials involved that there should be unification of the activities of agencies involved in earthquake prediction, possibly under the umbrella of an "earthquake agency".

But, according to Shimamura, apart from agreeing to pool data at the Japan Meteorological Agency (JMA), all discussion of unification was "killed" by Harumi Aoki, chairman of the subcommittee, on the grounds that it was "too early" to discuss such matters and there was "no time".

Following the Kobe disaster, government agencies and ministries involved in the prediction programme, which employs about 500 researchers on an annual budget of more than ¥10 billion (US\$120 million), called for an unscheduled review of the programme in the hope of increasing its budget (see *Nature* 374, 205; 1995).

In a departure from present practice, the report of the review committee calls on the JMA to act as a central agency for the collection of all data on earthquakes, including microearthquakes. Until now, microearthquakes have been the exclusive domain of university researchers.

The report also calls for more research on active faults, such as the one that caused the Kobe earthquake. In other respects, its conclusions are similar to a lengthy review of the programme carried out in 1992 and 1993, which called for only minor adjustments in the 30-year-old programme (see *Nature* 364, 370; 1993).

But in addition to Shimamura, some members of the subcommittee are also upset by the way the recommendations were prepared. A draft was drawn up by

Tomowo Hirasawa of Tohoku University, one of the senior leaders of the programme, with a small drafting committee.

The committee included Aoki and a few select members of the 20-man earthquake prediction subcommittee. But despite "heated discussions" at the subcommittee's final meeting, changes to the draft were only "cosmetic" says Shimamura.

Aoki dismisses Shimamura's criticisms as "emotional" and "selfish", while Yukiko Hirakawa of the Ministry of Education, who has administrative responsibility for the programme, says she cannot believe a member of the subcommittee made such criticisms.

Aoki claims to have made great personal efforts to explain earthquake prediction to the general public, for example through the publication in 1990 of a pamphlet entitled "*Earthquake Prediction Now*". He says that both critics of the programme and young researchers have failed to come forward with concrete proposals for reform.

But Shimamura is not alone in his criticisms. Seiya Uyeda, a professor of geotectonics at both Tokai University and Texas A&M University in the United States, and a member of the Geodetic Council, wrote to the minister of education during the review process stressing the need for a more thorough review. In particular, he urged the substantial involvement of scientists outside the programme covering a wide spectrum of Earth sciences.

Yoshihide Kozai, emeritus professor of the National Astronomical Observatory, and chairman of the Geodetic Council, who was responsible for submitting the recommendations to the Ministry of Education, says that he also sees the need for an external review of the programme.

Kozai says he has called for such a review on several occasions, but that it is not yet clear when it will be carried out, or what form it will take. Nevertheless, he feels that the present plan for the programme is better than the previous one. In particular, he points to the decision to collect all data centrally at the JMA as a significant step forward.

There is considerable disagreement, however, among members of the subcommittee about what the decision regarding JMA really means. Some see it as a move to reduce the heavy and monotonous work of collecting data on 'microquakes' at universities. Others are determined that universities should maintain a central role in monitoring such earthquakes.

David Swinbanks

Protein institute gets a reprieve — and a new role

Tokyo. Japan's generously funded Protein Engineering Research Institute (PERI) in Osaka is to get a second lease of life, under the new name — and expanded scientific goals — of the Biomolecular Engineering Research Institute (BERI).

PERI is one of a new breed of semi-private institutes jointly funded by private industry and the Japan Key Technology Center — known as Japan Key-TEC — a semi-governmental organization supported by dividends of government-held shares in the telecommunications company Nippon Telegraph and Telephone Corporation, which was privatized in 1985.

These institutes provide an unusual research environment for Japan. Scientists are given freedom to carry out basic research — much as they would be in a university — but are also provided with exceptionally good facilities that are comparable to, if not better than, those in private industry.

PERI is a small institute with about 60 researchers, drawn from universities and industry. Since its establishment in 1986, it has received more than ¥17 billion (US\$210 million), 70 per cent coming from Japan Key-TEC, and the rest from 14 companies, including Kyowa Hakko and Takeda Chemical Industries.

The institute has established a relatively strong reputation in protein research. But it was due to close down next March — ten years after it was founded — under the rules of the Japan Key-TEC system which limit the life of institutes to ensure that they do not become entrenched in a particular line of research.

PERI will now be reborn as BERI, and will receive another ¥15 billion in funds over 8 years, provided jointly by Japan Key-TEC and an expanded consortium of 18 companies. These include several newcomers — Japan Tobacco, Hitachi and Hoechst Japan, and the pharmaceutical companies Sankyo, Shionogi, Tanabe and Yamanouchi.

BERI was formally established as a corporation on 28 March. It is in the same building as PERI until the latter ceases to exist next year. At that time, Yoshiro Shimura, a biophysicist from Kyoto University, will take over as head of the new institute, and about half of 20 leading researchers of the former PERI will be replaced.

PERI has concentrated its research efforts on determining the structure of proteins. BERI will expand this goal to look in addition at lipids, sugars and nucleic acids; it will also attempt to establish their biological function as well as their structure.

D.S.

Chile promises to try harder over VLT dispute

Garching. A dispute over the ownership of land in Chile on which the European Southern Observatory (ESO) is building the world's largest optical telescope has come a step closer to resolution with a promise by the Chilean government to take steps to protect ESO's right to continue its work unhindered.

The five-year dispute has delayed construction work in preparation for the DM570 million (US\$800 million) Very Large Telescope (VLT), on Mount Paranal in northern Chile, and has led ESO to consider moving the operation to another country (see *Nature* 370, 494; 1994).

Last week, the government issued a statement promising to seek a clear ruling from the country's Senate on ESO's immunity from national laws. Riccardo Giacconi, ESO's director general, has welcomed the move, which came after a week of hard negotiation between himself, the foreign ministry and the president of Chile.

But he adds that ESO will not abandon its studies of ways of pulling out of Chile for at least another two months, during which it should become clear whether the government's promise will be fulfilled.

At the same time as announcing the government's commitment to protect construction of the VLT, José Miguel Insulza, the Chilean foreign minister, signed with Giacconi a supplementary agreement to ESO's original 1963 treaty with Chile.

Drawn up after complaints of unfair treatment by Chilean scientists and workers, the new agreement includes two concessions made by ESO. The agreement guarantees privileged viewing time on all ESO telescopes — including the VLT — for Chilean astronomers, and that ESO will respect national labour laws which, for example, confer rights to free collective bargaining.

The agreement reiterates ESO's immunity from national law and specifies that this immunity applies to all ESO sites in Chile. This is the key issue in the affair, as ESO's immunity has not been observed by the judicial system. The courts have not yet decided whether they accept the immunity agreement signed by the government.

The VLT site around Mount Paranal in northern Chile was donated by the Chilean government to ESO in 1988 in the belief that the site constituted public land. But two years later a powerful local family, the Latorres, claiming ancestral ownership of the land, sought compensation from the government through the courts.

Several injunctions to stop construction work at Mount Paranal have been issued at regular intervals over the last year, the latest one in January. The situation came to a head last month when a court official, under police escort, entered ESO premises to take note of work carried out in defiance of the

courts (see *Nature* 374, 484; 1995).

Citing its immunity from local legal disputes, ESO has — after complying for three weeks with the first injunction — ignored further court orders to stop work. The response to the first injunction "was intended as a gesture of good will", says Giacconi. It will never be repeated, he adds.

ESO estimates that the delays could add between DM20 million and DM50 million to the final VLT bill, largely as a result of the Latorre family's influence. ESO officials cite

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Very Large Telescope: delays could increase construction costs by up to \$36 million.

numerous occasions when building materials have been held up in customs, and this and other delays has made it impossible for ESO to keep to schedule.

ESO has long complained that the Chilean government has taken no practical steps to defend its immunity, despite continual verbal reassurances. The root of the government's hesitancy is a power struggle between the government and the judicial system, which are constitutionally separate.

According to Rodrigo de Castro, an ESO

advisor based in Chile, the government has faced two options: to plead ESO's case to the Supreme Court through the state defense council, an group of lawyers nominated by the government to defend the Chilean state, or to ask the Senate to rule on which body should have precedence.

But so far it has avoided such confrontational measures, afraid of inflaming the dispute over the issue. "Until now the government had always been very vague on its intentions," says de Castro. "But now it has said that it will ask for a Senate ruling. This is something quite new for us."

ESO remains cautious. At an extraordinary meeting of the ESO council last week, delegates from the organization's eight member states agreed to give the Chilean government two months to demonstrate that it genuinely intends to resolve the land ownership dispute.

Delegates also agreed to set up a formal working group to investigate other possible sites for the telescope if Chile fails to deliver on its promise. One alternative is a site in Namibia owned by Germany's Max Planck Society. The working group will also consider potential sites in the northern hemisphere, and will report back to June's council meeting, which will decide whether the VLT should stay or move.

But there are strong grounds for optimism that the telescope, which consists of four, eight-metre optical dishes and three smaller telescopes, will still be built in Chile. Already, there have been two indications that the government is taking its new promise seriously. First, a second police visit to the construction site — threatened by the local court — did not materialize.

Furthermore, a court order issued in April, instructing customs officials not to release technical and building materials for the VLT, has been lifted. **Allison Abbott**

India seeks source of plague 'peculiarities'

New Delhi. The Indian government, concerned that last year's outbreak of plague in Surat may have been caused by a new strain or mutant of *Yersinia pestis*, is seeking the help of foreign microbiological laboratories in precisely identifying the culprit organism with the use of DNA fingerprinting techniques.

Government officials say the analysis will take up to two months to complete. As a result, a report on the outbreak by a committee of experts, previously due to be submitted by 30 April, is expected to be delayed by three months.

In its preliminary report, the committee, headed by Vulimiri Ramalingaswami, confirmed that the outbreak was indeed plague, on the basis of *Y. pestis* isolated from two cultures (see *Nature* 373,

650; 1995). To confirm these findings, the health minister now wants the strain to be identified by at least three international laboratories.

According to health ministry officials, detailed identification of the strain is essential because of peculiarities about the epidemiology of the Surat outbreak. Only about 20 deaths in Surat could be attributed to pneumonic plague, and — despite being highly infectious — the disease did not spread outside the town.

"The genetic typing will throw light of whether Surat was struck by a non-virulent mutant," says U.V.S. Rao, a microbiologist with the National Institute of Communicable Diseases in Delhi. "If that is found to be the case, we want to know where this strain came from." K.S. Jayaraman

France may break ranks over space station

Paris. A French government report on space science last week recommended that France proceed independently of the European Space Agency (ESA) and develop a crew rescue vehicle (CRV) for the international space station Alpha in direct cooperation with Russia.

The CRV was dropped from ESA's planned contribution to the space station last month to reduce the plan's overall costs from ECU3.5 billion (US\$4.6 billion) to ECU1.8 billion (see *Nature* 374, 586; 1995).

ESA's new plans now consist merely of the Columbus Orbiting Facility (COF), which will be built by Germany and Italy, and the Automated Transfer Vehicle (ATV), a relatively low-cost craft that will ferry freight and fuel to the station, and will be built in France.

Last week's report, which was prepared after discussion with aerospace scientists and industrial companies by François Fillon, the minister of research and higher education, recommends that France should now build the CRV with the Russian space agency, RKA.

The craft would differ from the original design in that it would now be based on the existing Russian Soyuz vehicle, upgraded with technology developed in the CRV programme. The new vehicle would be compatible with the launcher Ariane V, as well as with the space station.

But such a scheme would make it more difficult for ESA to find the money for its own space station plans. The agency has so far obtained commitments from member states to cover just ECU1.2 billion of the costs of its scaled-back space station plans. Fillon's report recommends that France should not increase its current minimal contribution, namely 10 per cent of the costs of COF and 25 per cent of the ATV.

Fillon's report, entitled *Refinding a Long-Term Vision for Space Science*, was submitted last week to Edouard Balladur, the prime minister. Balladur was defeated in a bid for the presidency in the first round of France's elections last Sunday (23 April). But many observers feel that the plan to upgrade Soyuz is likely to win political and budgetary support, irrespective of who is elected president on 7 May.

The proposal comes at a time when France's National Centre for Space Studies (CNES) is already cutting back on its contribution to ESA, diverting the money saved towards either national programmes or bilateral and multilateral collaboration outside the agency. The proportion of CNES's budget allocated to ESA has decreased from more than half to just 40 per cent this year as a result of its withdrawal from elements of ESA's 'optional' programmes, such as those associated with the space station.

French officials say that their disenchant-

ment with ESA is due, in particular, to the deadlock caused by the agency's budgetary problems — which have been caused by a prolonged economic recession in member states — and continuing uncertainty over the future of its manned programmes.

Some feel that France, still the largest contributor to ESA's budget, wants to use its muscle as the world's third largest space power to plough its own furrow more effectively. But Aerospatiale, the country's main aerospace manufacturer, argues that this need not mean that France is becoming "less of a good European".

Company officials point out that European aerospace companies already need to work as consortia to be internationally competitive, and that the government wants only to "put industry in a position where it can defend itself in negotiations with other European companies."

But Mark Giget, of the Paris-based space consultancy Euroconsult, argues that Fillon's report, widely seen as reflecting a

strong current of opinion in government circles, is "not very European".

Giget argues that the report represents a missed opportunity for France to have proposed the bold and firm commitments that Europe needs if it is to win long-term support for space from its member states. Although it does recommend that ESA produce proposals for both a moon mission and a specific Earth observation programme by 1997, he argues that the details are vague, and focus too much on French interests.

In any case, much of the report could still end up gathering dust on a ministry shelf as a result of the election. Fillon — an active supporter of Balladur — is likely to be removed from the government. Jacques Chirac, the Gaullist candidate widely tipped as the eventual winner, has already promised that, if elected, he will carry out his own review of the French space programme before the meeting of Europe's space ministers in Toulouse in October.

Declan Butler

Concorsi 'corruption' leads to court

Munich. In the first case of its kind, eight Italian professors have been suspended from their academic duties after being charged with corruption in making academic appointments.

The eight are accused of unfairly appointing their young collaborators (*allievi*) in preference to better qualified candidates in two national competitions, known as *concorsi*, to appoint a total of 25 professors in the medical discipline ENT (ear, nose and throat).

All eight, who have been barred from academic activity pending the outcome of a court case, face criminal charges. Allegations of nepotism are associated with at least two of the appointments.

Seven of the accused were among the ten members of the committees appointed by the research ministry to run *concorsi* in 1988 and 1992. The eighth is accused of making a deal with committee members to ensure the appointment of his son.

The results of *concorsi* have occasionally been overturned in the past on procedural grounds. But this is the first time that a criminal action has been brought in connection with the system under which a *concorso* is held every few years by the ministry of universities and research to arrange appointments to full-time position.

Shortly before the competition takes place, universities throughout Italy must apply to the research minister for new posts in different disciplines. The minister then appoints committees for each discipline, which both run individual

concorsi and choose the winners.

Once the winners have been chosen, each is allocated to a university with a vacant position.

Critics of the system complain that appointments too often involve deals between powerful departmental heads, irrespective of whether they sit on the committees, and that the whole procedure lacks both transparency and clear criteria for appointments to academic positions.

So far, attempts by successive research ministers to introduce such measures have failed because of lack of support in parliament, where many members are themselves professors. But last week, Italy's current research minister, Giorgio Salvini, won government approval for a bill which modifies the *concorsi* rules.

If the bill is approved by parliament, *concorso* committees in future will be required to select more candidates than the number of positions available. A shortlist of candidates would thus be prepared, from which individual universities could then make appointments.

Salvini wants the bill to be approved rapidly so he can call the next *concorso* as soon as possible under new rules. The bill has been granted accelerated passage through parliament to facilitate this. It is likely to be approved because it is in line with new rules on the autonomy of universities (see *Nature* 365, 682; 1993), and neatly sidesteps the thornier issue of criteria for minimum qualifications — a fact that has not gone unnoticed by Italian academics.

Alison Abbott

Panel to gather intelligence on rival technology efforts

Paris. Luc Montagnier, of the Institut Pasteur in Paris, may need to swap his lab coat for a trench coat after his surprise appointment to a committee set up by the French government to improve the collection and distribution of 'economic intelligence'.

The Committee for Competitiveness and Economic Security, as it is known, was set up earlier this month by presidential decree. Its seven members will be appointed to serve a three-year term of office by the prime minister, who will himself chair the committee.

In a report to President François Mitterrand accompanying the decree, Edouard Balladur, the current prime minister, claimed that changes in the geopolitical scene — particularly the globalization of business and technology — have made the committee necessary.

The committee's creation was prompted by a report on business intelligence by the Commissariat Général au Plan, a powerful agency made up of industrialists and representatives of trades unions and other public organizations. It concluded that the collection and distribution of economic intelligence had suffered from "multiple walls" and a "lack of coordination".

France has been spying on competitors' technologies and business strategies for years. But the commissariat argued that countries such as the United States, Germany, and Japan had much more advanced systems of 'legal' intelligence.

The United States has traditionally left the gathering of economic intelligence pri-

marily to large companies. But the administration of President Bill Clinton has also begun to seek ways of improving the flow of information to those companies that might most benefit from it, even though — unlike Japan, for example — it has yet to use economic intelligence to support a national industrial policy.

The French committee appears similar to Japan's Cabinet Information and Research Office, and will be based within the General Secretariat of the National Defence (SGDN), which analyses intelligence, including technology developments, for the prime minister. Remy Pautrat, SGDN's deputy-secretary general, is a former head of the French counter-espionage agency DST.

Apart from Montagnier, the committee includes Bernard Esambert, a financier who is also president of the board of the Institut Pasteur, Philippe Jaffré, the chairman of Elf Aquitaine, and Henri Martre, honorary president of Aerospatiale and chairman of the working group that prepared the commissariat's report.

The creation of the committee has been generally welcomed by observers. But some also suspect a political motive. The Paris-based publication *Intelligence Newsletter*, for example, points out that many members of the committee are close to Balladur, who nominated them just before he was eliminated from the presidential elections last Sunday. The apparent political dimension of some of these appointments may limit the committee's effectiveness. **Declan Butler**

US physicists seek construction role in Europe's LHC

Washington. An informal consortium of three US laboratories is working out how the United States could contribute to the construction of the proposed Large Hadron Collider (LHC) at the European Laboratory for Particle Physics (CERN) in Geneva.

Experts from the Brookhaven National Laboratory, Fermilab and the Lawrence Berkeley Laboratory have started talks with CERN about a material US contribution to the accelerator project which could be worth between \$200 million and \$400 million, according to Peter Limon of Fermilab.

The contribution would probably be made in the form of engineered components, such as superconducting magnets. US high-energy physicists argue that such involvement in the construction of the LHC is essential if the United States is to be fully engaged in the project.

US physicists are already involved in the design of two detectors to be installed at the LHC — Atlas and CMS — but want their involvement to go further. "No-one has said we can't be part of the detectors if we aren't part of building the machine," Limon last week told a meeting of the American Physical Society in Washington. But without involvement in the construction "it would be the last time" that the United States would be invited to share in the scientific output of such an international collaboration.

The physics community and its paymaster, the Department of Energy, are united behind such involvement, according to Sidney Drell, deputy director of the Stanford Linear Accelerator Center in California. Last year, Drell prepared a report for the US government on the future of high-energy physics. This suggested that LHC could be given up to \$400 million in US support without hurting domestic programmes. But he admits uncertainty over whether Congress, which will ultimately make the decision, would be so supportive.

The Department of Energy has asked for relatively small budgets of around \$10 million for each of the next two financial years to support work related to the LHC. Only in the year after that — 1998 — would a more substantial amount be needed to support actual LHC construction. So Congress need not decide on the matter for two years.

According to Limon, one option being considered for US participation would involve supplying all the quadrupole magnets used to focus the particle beam. "No dollar bills would actually go to CERN," he said. "The money stays here and the goods go there". He claimed that Japan is considering contributing up to \$100 million for the construction of the accelerator, and \$50 million for its detectors. **Colin Macilwain**

Glaxo researchers move to new home

London. **Glaxo Wellcome, the world's largest pharmaceutical company after the recent merger between Glaxo and the Wellcome Foundation, last week opened a £700 million (US\$1.1 billion) medicines research centre in Stevenage (right), north of London.**

The centre has been designed to bring together drug discovery teams from a range of disciplines, including organic and analytical chemistry, microbiology, bioinformatics, genomics, pharmacology and cellular and molecular biology.

By the end of June, Glaxo Wellcome will have moved 1500 scientists and support staff on to the site, the completion of a project which began

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

seven years ago. "Stevenage is one part of our overall strategy designed to enhance our R&D productivity," says one leading Glaxo research official.

The centre has been designed to encourage researchers in different disciplines to interact, for example by using numerous meeting places — known as 'nodes' — on the main thoroughfares between buildings. □

Glaxo plc

Hong Kong inquiry dispute

SIR — We wish to comment on the letter from Mr John Griffiths on matters arising from our legal dispute with Dr T. H. Lam and the subsequent inquiry by the University of Hong Kong¹. The Hong Kong University inquiry consisted of 14 closed meetings during 16 months. We were excluded from being a party. Griffiths, the barrister defending Lam, claims that new evidence was presented at the inquiry. We have not been allowed to see it.

Griffiths' comment that the university inquiry "received far more evidence than was called at the civil trial" is difficult to understand as the high court trial lasted at least three times longer and went over more than 20,000 pages of documents in open court. As stated in the judgement from the court of appeal², "The trial lasted some 9 weeks and every conceivable point which could be raised was... it is clear the Judge took the greatest care in analysing the evidence... I am in no doubt that not only has derivation been proved, but also Dr Lam has copied a substantial part of Dr Koo's questionnaire. He has used someone else's skill and labour and unfairly produced it as though it was his own."

Griffiths cites the comments of 26 epidemiologists, but fails to point out that they received insufficient documents for analysis. There were 14 versions of our questionnaire and 4 of Lam's in dispute. A comparison of these documents with 8 other epidemiological questionnaires was made during the high court trial but not the university inquiry.

Griffiths is also inaccurate in saying that the sworn evidence of Koo at the high court trial was contradicted by the statements of two research assistants in the university inquiry. In the high court judgement³, it clearly states that "The 'used' Koo-Ho questionnaires are four in number... KH4a was used by Miss Carol Tong... KH4b was used by Miss Nancy Lee... KH3c is a translation into Chinese of KH4a... KH4d is the version kept by Dr Koo herself." Who possessed and used which version was not disputed by Griffiths in the court of appeal and Lee's testimony in the university inquiry did not contradict the high court judgement.

Griffiths' comment that our questionnaire was "inferior" is contrary to published assessments. From the 1986 US Surgeon General's report on *The Health Consequences of Involuntary Smoking*⁴ it is said: "The design of this study addressed the criticisms of other studies that an index of involuntary smoking exposures based only on spouses' smoking habits is inadequate, and broadened the exposure assessment to include all locations of tobacco smoke exposure." Others have said "I think we should encourage original research like that of Freedman 1983 on

surrogates for ETS and Koo's 1989 paper..."⁵.

Griffiths and the university inquiry criticize the effects of the court judgements on the "free exchange of ideas" between scientists. Voluntary exchange of information between scientists is one thing, but when a colleague surreptitiously obtained 2 questionnaires that we laboured to produce, used it for his benefit and to our detriment, and without acknowledgement, then the issues at stake are morals and ethics.

J. H.-C. Ho

Department of Radiation Oncology,

L. C. Koo

Department of Community Medicine,

University of Hong Kong, Hong Kong

1. Griffiths, J. *Nature* **374**, 301 (1995).
2. *Hong Kong Law Reports* **1**, 329–352 (1994).
3. *Intellectual Property Reports* **23**, 607–640 (1993).
4. US Surgeon General, *The Health Consequences of Involuntary Smoking*, 81 (US Dept of Health and Human Services, Washington, DC, 1986).
5. *Environmental Tobacco Smoke 127* (eds Ecobichon, D. J. & Wu, J. M.) (Lexington Books, Lexington, Massachusetts).

Brazilian science

SIR — I would like to comment on the report by your São Paulo correspondent on science in Brazil (*Nature* **373**, 274; 1995) which contains some statements that are misleading and incorrect. First, it is obvious to Brazilian scientists that a measure of "reprieve from the hard times" they have faced came only after the fall of the government of Fernando Collor de Mello in 1992, and not since the beginning of the Collor administration, as your report states.

Second, the new president, Fernando Henrique Cardoso, is a well known social scientist who has since October 1992 been a minister first of foreign affairs, and subsequently of finance. He was therefore well aware of the importance of science and technology, and was an important participant in the reprieve of science and technology during Itamar Franco's administration.

The retention of the minister for science and technology might therefore be attributed to an appreciation of the minister's past performance both by the scientific community and by the new president, to whom he has been known for more than 40 years, rather than to a "low political significance" being attributed to the sector, as stated in your report.

Third, I have not made any claims, as your report suggests, to the rescue of the Programme of Support for Scientific and Technology Development (PADCT) or for obtaining funds for university-based research groups. Nevertheless, it remains

true that there was an overall reduction of 50 per cent in the resources allocated to science and technology between the start of the Collor administration in 1990 and October 1992. Between then and 1994, we have not only returned to the pre-Collor level of spending, but have also obtained resources arising from privatization policies which have allowed for the reactivation, among other projects, of a certain number of university-based research groups.

José Israel Vargas

Ministerio de Ciencia e Tecnologia,

Esplanada dos Ministerios,

Bloco E, 4º Andar,

70067-900 Brasilia,

Brasil

■ The apparent claim that Brazilian scientists have faced hard times since the fall of the Collor administration was a mistake introduced during editing. The other two issues raised by the minister are open to varying interpretations. Vargas's claim to have helped rescue the PADCT was made in an interview last year with the article's author. □

Future biodiversity

SIR — Arguments for the ethical incentive to sustain biodiversity for future generations cannot be simply dismissed by analogy to "finite" abiotic resources such as water and air¹, or lead². Elements are generally being moved around by humanity, not destroyed. Future generations could mine our wastegrounds for discarded 'non-renewable' resources, and if necessary use energy for restoration of molecules from component elements. Further, substitute energy sources can be anticipated (there is a lot of energy in the Solar System). In contrast, very few species could be re-created³: you cannot re-create what you never knew, and our descendants will not even know how many species we extirpated⁴. Biodiversity is being irreversibly, consciously, and often unnecessarily destroyed; Beckerman's treatment of this central issue is superficial¹. Our failure to promote the survival of as many species (or rather their relatively immortal genes⁵) into the orbit of future generations is less ethical than theft.

Clive Hamblen

University of Oxford,

Department of Zoology,

South Parks Road,

Oxford OX1 3PS, UK

1. Beckerman, W. *Small is Stupid* (Duckworth, London, 1995).
2. Maddox, J. *Nature* **374**, 305 (1995).
3. Hamblen, C. *Report on the Environmental Consequences of Uncontrolled Population Growth* (Council of Europe, Strasbourg, 1994).
4. May, R. M. *Science* **241**, 1441–1443 (1988).
5. Dawkins, R. *The Selfish Gene* (Oxford Univ. Press, Oxford, 1976).

Beyond Einstein's theory of gravitation?

The possibility that the law of gravitation was different early in the history of the Universe seems to be coming into fashion again — and it may just be measurable.

We all know that Einstein's theory of gravitation is one of the most remarkable and successful theories ever. It is remarkable because there is a sense in which it is the product of pure thought. The starting point is the Equivalence Principle, which goes back to Galileo: there is no way of distinguishing by means of strictly local observations between the circumstances in which an object is accelerated by means of an external gravitational field (as when an apple falls to the Earth) and those in which it is observed from a uniformly accelerated frame of reference.

The identity of the gravitational and inertial masses of an object, often offered as a consequence of the Equivalence Principle, is logically independent. Why, otherwise, would Eötvös in the 1920s have gone to such trouble to measure the gravitational forces on objects of different chemical constitution? And why, otherwise, would there have been such excitement, just a decade ago, in the idea of the "fifth force"?

These were two of the cornerstones of Einstein's theory. He also, of course, knew that any theory of gravitation must merge into Newton's law of gravitation when the masses are small, and that any kinematical implications of the theory must approximate to his own special theory of relativity in appropriate circumstances. His objective was to represent geometrically an arbitrary gravitational field; the outcome was a set of tensor quantities representing the geometry of space-time, with rules for telling where the masses are buried. That is the sense in which the theory is pure thought. It will always be revered as such.

It has also been remarkably successful. The bending of light near the limb of the Sun at the solar eclipse of 1919 was the first validation of Einstein's theory of gravitation. The precession of the perihelion of Mercury came later. The notion that a gravitational field will affect the frequency of a spectral line cannot easily be verified astrophysically, but the prediction has been confirmed by measurements of Mössbauer spectra in the Earth's gravitational field. As it happens, there are few other tests of general relativity than these (but the precession of perihelia now extends to pulsars in elliptical orbit about a companion star).

Whether or not this is a sufficiently rich harvest of validation is a matter of taste, but there is little doubt that Einstein's theory of gravitation has become revered (as it is) because it also accommodates a nat-

ural explanation of the expanding Universe. De Sitter's first solution of Einstein's equations expanded spontaneously, but was devoid of matter. Luckily, Friedmann in Moscow showed that there is a whole range of solutions of these difficult equations that will accommodate matter (provided that it is distributed homogeneously and isotropically). These are the solutions on which all present models of the Universe are based.

To be sure, there are a few clouds on the horizon, but they seem hardly "bigger than a man's hand". One difficulty is that general relativity allows black holes to exist, with all the disbelief that that requires of the rest of us. The most serious difficulty is that, after more than a quarter of a century, there is still no way of reconciling Einstein's theory of gravitation, his general theory of relativity, with quantum physics. A generation of talented people, not so much an army as a company, has beaten its head against that problem without much success. Some have begun wondering whether there is some other way, perhaps a theory of gravitation that is easily reconcilable with quantum mechanics or, otherwise, does not allow the dense concentrations of matter that make quantum mechanics necessary.

Now there is fresh hope that they may have nature on their side, and even that their optimism may be verifiable. Nostalgia enters at this point. A quarter of a century ago, there was the Brans-Dicke alternative to Einstein's theory. Following the same intellectual construction as Einstein's, it included a scalar as well as a tensor field in its description of space-time. It was eventually undermined by the recognition that the scalar field would now be much more conspicuous than observations of the real world will allow. Unfortunately for Brans and Dicke, their theory was decently put to rest before Alan Guth invented the notion that, in the very earliest moments of the evolution of the Universe, a few microseconds after the Big Bang, the condensation of radiation into matter caused a huge inflation of the physical size of the Universe (by a factor of 10^{30} or thereabouts).

Two years ago, Thibault Damour and Kenneth Nordvedt from the Institut des Hautes Etudes at Bures sur Yvette (but with separate affiliations to the Observatoire de Paris and the University of Montana respectively) looked again at scalar-tensor theories of gravitation and

concluded that inflation would upset earlier conclusions (*Phys. Rev. Lett.* **70**, 2217-2219; 1993). In particular, they showed that, in an evolving Universe, the scalar part of a scalar-tensor field would rapidly decay as a consequence of its inherent structure. No traces of the scalar field would be left at the present time. But a bout of inflation could cure that, leaving a barely measurable curvature of space-time as a proof that the space-time metric had once had a scalar component.

Benjamin Lange, from the Department of Physics at Stanford University, believes he has a way of measuring this curvature. He wants to put into an orbit about the Earth a gyroscope whose stability would far outdo the (mostly) *gedanken* experiments of the late William Fairbank of the same university. (*Phys. Rev. Lett.* **74**, 1904-1907; 1995). It is an astonishing device: a 5-cm sphere of doped silicon spinning in empty space without support.

Naturally, even in a high orbit, a spinning sphere meant to act as a gyroscope must be protected from drag; why not enclose it in a spherical container, or satellite, which can also conveniently carry the equipment meant to read out data from the spinning sphere? The obvious objection that the casing (exposed to drag) and the sphere (free from it) would then collide is met by the simple device of measuring where the sphere is and arranging that the satellite and its equipment keep their designed distance.

That, of course, is almost child's stuff now. The professional cleverness in Lange's scheme is to use a set of Helmholtz coils to spin up the gyroscope and to collect data from the system. In a helpful letter, Lange explains that he hopes to measure the orientation of the axis of the spinning sphere to as little as 0.04 microseconds of arc by means of collimating instruments looking at optically flat surfaces at the two poles of the sphere. To cancel out the proper motion of the star on which the telescope is aligned, it would be better to have two counter-rotating satellites in orbit at the same time.

And the pay-off? If there were ever a scalar field, its remnant might just be measurable. Lange argues that his satellite gyroscope is more likely than any other device, by an order of magnitude, to tell whether gravitation now is different from what it used to be in the early Universe. And that would be a prize worth having.

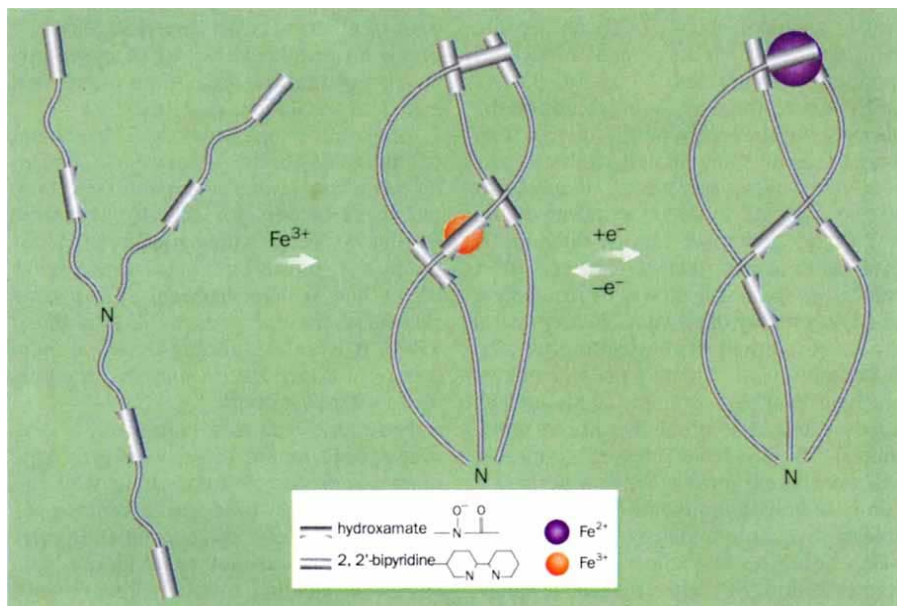
John Maddox

Iron flicks the switch

Edwin C. Constable

EVER since Kekulé's dream of benzene as a snake grasping its tail in its mouth, chemists have been describing molecules in allegorical terms. Originally, this was driven by the convenience of dealing with concepts relating to familiar macroscopic objects, but increasingly chemists are now seeking to reproduce the structure and function of these objects at the molecular level. Currently, many research teams are motivated by the vision of molecular or supramolecular machines

formation of an orange compound with a triple-helical structure in which the three hydroxamate groups of the strands are coordinated to an approximately octahedral iron(III) centre. In other words, the iron(III) has selectively found the binding site preferred by higher-oxidation-state metal ions, and in so doing has twisted the three strands into a triple-helical topology. A similar reaction with iron(II), in contrast, yields a purple compound, also with a triple-helical structure,



Molecular switch¹ — on coordination to iron, the three-stranded ligand twists into a helix with two possible binding sites. Iron(III) prefers the hydroxamate site, but if it is reduced to iron(II), it flicks into the 2,2'-bipyridine site.

and molecule-based technology. Machines, of course, need moving parts, and electronic circuits need switches. On page 790 of this issue¹, Shanzer and co-workers describe an elegant approach to molecular switches.

Their starting point is a molecule in which a central nitrogen atom has three strands attached to it. Each of these strands contains two metal-binding domains; one of these is a 2,2'-bipyridine group and the other is a hydroxamic acid ($-N(OH)CO-$) functionality. The neutral 2,2'-bipyridine region binds strongly to low-oxidation-state ions, whereas the hydroxamic acid may be deprotonated to give a hydroxamate ($-N(O^-)CO-$) group which favours the binding of higher-oxidation-state species.

Although randomly oriented in the free ligand, these chains are pre-organized for the formation of intertwined helical structures upon coordination to a metal centre (see figure). Reaction of these ligands with iron(III) salts results in the

in which the metal is bound to the 2,2'-bipyridine region. So if the iron(II) and iron(III) complexes are interconverted upon oxidation or reduction respectively, the metal ion must flick across to the site appropriate for the other species. The difference in colour between the iron(II) and iron(III) states means that it is easy to tell whether the switch is 'on' or 'off'. In essence, Shanzer and co-workers have shown that it is possible to build, address and read a molecular switch, operated by the gain or loss of an electron.

Aspects of this work fit into some developing trends in supramolecular chemistry. Although the use of metal ions to control the assembly of helical structures is well known², few examples have been found in which discrete metal-binding sites are coded for particular types of metal ions^{3–5}. The new work builds on earlier studies of systems containing only a single type of metal-binding domain⁶. One feature of the present work that offers interesting opportunities for the

future is the introduction of chiral chains, using amino-acid links. When chiral chains with L-amino acids are used, a single diastereomer with a left-handed configuration at the chiral (D_3 symmetry) metal centre is formed. The methodology introduced by Shanzer and colleagues opens prospects for the design not only of switches but of metal-ion selective switches and possibly chirally selective switches.

The ligands used by Shanzer and co-workers are fairly pre-organized, in the sense that the three chains are covalently linked. It is also possible to use the metal ion as an assembly principle to bring together a number of helical subunits — a process of obvious relevance to biological self-assembly. A nice example, which illustrates the power of selective metal-ion-based assembly, was reported a few years back by Minoura⁷ who designed a crown ether with a single α -helical polypeptide chain attached to it. The crown ether chosen, benzo-15-crown-5, is the correct size for a sodium ion to fit inside the cavity, but it is too small for the larger potassium ion. As a consequence, interaction with sodium ions gives a monohelical species of 1:1 stoichiometry, whereas interaction with potassium ions gives a system in which two crown ether molecules 'sandwich' the metal ion. In this case at least, Minoura showed that helix aggregation can be metal-ion specific.

The next step is to introduce helix-helix interactions within the aggregates, and a series of papers over the past six years described the metal-controlled assembly of two⁸, three^{9–11} and four-helix¹² bundles. The principle behind all of these studies is similar, with α -helical polypeptide chains being covalently linked to metal-binding domains, and the helix bundle is formed on coordination to appropriate metal centres. For instance, the two-helix bundle may be formed by the attachment of two polypeptide chains to a 2,2'-bipyridine metal-binding domain, followed by coordination to a six-

1. Zelikovich, L., Libman, J. & Shanzer, A. *Nature* **374**, 790–792 (1995).
2. Constable, E. C. *Tetrahedron* **48**, 10013–10059 (1992).
3. Piguet, C., Hopfgartner, G., Bocquet, B., Schaad, O. & Williams, A. F. *J. Am. chem. Soc.* **116**, 9092–9102 (1994).
4. Piguet, C., Hopfgartner, G., Williams, A. F. & Bünzli, J.-C. *G. J. chem. Soc., chem. Commun.*, 491–492 (1995).
5. Constable, E. C., Edwards, A. J., Raithby, P. R. & Walker, J. V. *Angew. Chem. int. Edn Engl.* **32**, 1465–1467 (1993).
6. Libman, J., Tor, Y. & Shanzer, A. *J. Am. chem. Soc.* **109**, 5880–5881 (1987).
7. Minoura, N. *J. chem. Soc., chem. Commun.*, 196–197 (1993).
8. Mihara, H. *et al. Chem. Lett.* 1813–1816 (1992).
9. Lieberman, M. & Sasaki, T. *J. Am. chem. Soc.* **113**, 1470–1471 (1990).
10. Ghadiri, M. R., Soares, C. & Choi, C. *J. Am. chem. Soc.* **114**, 825–831 (1992).
11. Ghadiri, M. R. & Case, M. A. *Angew. Chem. int. Edn Engl.* **32**, 1594–1597 (1993).
12. Ghadiri, M. R., Soares, C. & Choi, C. *J. Am. chem. Soc.* **114**, 4000–4002 (1992).

coordinate ruthenium(II) centre in which four of the sites are 'blocked' by unsubstituted 2,2'-bipyridine ligands⁸. Actually, this work was rather more sophisticated, in that the polypeptide chains terminated in redox-active viologen groups, and photoactive anthraquinone substituents were attached between the metal and the viologen. The length of the polypeptide chain was such that the whole assembly functioned as a membrane-spanning device for photoconversion.

The interplay between metal ions and highly functionalized organic and biological ligands is not just a new playground for the supramolecular chemist. It holds very real promise of functional molecular devices in the next generation of technology. □

Edwin C. Constable is at the Institut für Anorganische Chemie der Universität Basel, Spitalstrasse 51, CH-4056 Basel, Switzerland.

VERTEBRATE ORIGINS

Conodonts join the club

Philippe Janvier

THE classification of the minute tooth-like or comb-like fossils known as conodonts, which existed from the Cambrian to the Triassic and are widely used by stratigraphers, has long been a bone of contention. The first description of an articulated conodont-bearing animal in 1983¹ offered persuasive arguments in favour of an affinity with vertebrates, but many specialists (including myself²) remained sceptical about its alleged vertebrate features.

Two papers on pages 798 and 800 of this issue now present fresh data on conodont structure and affinity^{3,4}. Gabbott *et al.*⁴ have new evidence for large eyeballs associated with eye muscles, and have discovered fossilized muscle fibres in the muscle blocks of the body; these fibres are strikingly similar to those that have been found in other, much younger, fossil fishes. Purnell³ provides evidence to support the long-disputed biting function of the conodont apparatus, in the form of wear facets on the tip of the cusps. He also concludes that conodonts were macrophagous and, consequently, that the earliest vertebrates could have been predators.

Considering this and other evidence for the vertebrate affinity of the conodonts that has accumulated over the past ten years^{5,6}, I think it is time for me to stop playing the Devil's advocate against this theory. There are still peculiarities in the structure of the 'conodont animal' that are at odds with what we know of classical early vertebrate fossils, but its tail, arrangement of the muscle blocks and large eyes are strikingly vertebrate-like.

Keeping in mind that extant craniates (animals with a skull and sensory capsules) include hagfishes and the vertebrates, and that the vertebrates include lampreys and the gnathostomes, the distribution of the characters observed in conodonts can be summarized as follows. The large notochord (not yet clearly documented) is a chordate character (shared by the tunicates, cephalochor-

dates and craniates); the chevron-shaped muscle blocks are shared by the cephalochordates and craniates alone; the radials in the caudal fin and the presence of eyes are craniate characters; the large eyes with extrinsic eye muscles are a vertebrate feature; the mineralized exoskeleton is shared with the Recent gnathostomes (jawed vertebrates) and with extinct, Palaeozoic armoured jawless vertebrates formerly referred to as 'ostracoderms'; the presence of bone cells (or enclosed dentine cells) in the dermal skeleton is unique to the gnathostomes and at least one 'ostracoderm' group, the osteostracans (see figure). This is all we can say from the available data.

Let us now enter the realm of interpretation. The organization of the conodont apparatus, namely the assemblage of minute tooth- or comb-like denticles (conodont elements) in the head of the animal, presents a bit of a problem. The recent reconstructions of some conodont assemblages⁷, with grasping ramiform ('S' and 'M') elements and transversally biting pectiniform ('P') elements, may fit a 'rasping tongue', more or less like that of the extant lampreys and hagfishes. This, however, would not necessarily mean that conodonts are close relatives of these extant groups, because current phylogenies imply that the 'rasping tongue' is a general craniate character, lost in the gnathostomes. Another possibility could be that the grasping S and M elements were attached to the lower lip, whereas the P elements were associated with a transversally biting structure derived from the velum of larval lampreys, which is thought to be the homologue of the jaws in jawed vertebrates.

The microstructure of the conodont elements is also ambiguous. Palaeontologists have recently tried to shoehorn the histological characters into the recognized classification of vertebrate hard tissues (enamel, dentine, bone, calcified cartilage)⁸, but none of the conodont hard tissues described so far unequivocally

matches those found in vertebrates. Although I now admit that the cell spaces of some conodont elements are strikingly similar to those of the dermal bone and mesodentine of vertebrates (a special type of dentine found in some early fossil vertebrates), I still regard conodont histology as unique to this group, being possibly derived from a more conventional type of early vertebrate histology.

Two features of the conodonts still puzzle me. One is the absence of any trace of a gill apparatus, be it in the form of imprints of gill pouches or traces of cartilaginous gill arches. These organs represent dense masses of organic matter and generally leave some trace in fossil vertebrates that are preserved in the same way as the conodont animals. The lack of a gill apparatus in conodonts may be explained by cutaneous respiration, but this sounds like an *ad hoc* explanation, as no other fish lacks gills. The other peculiar feature is the position of the eyeballs, apparently at the anterior tip of the head. This also occurs in some early jawless vertebrates,

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Possible interrelationships⁹ of the craniates ('ostracoderms' shown in brown; pteraspido-morphs including arandaspids, astraspids and heterostracans) and the two possible positions of the conodonts (1 and 2). Position 1 is supported by the mineralized exoskeleton alone, and position 2 by the exoskeleton and cellular dermal bone or mesodentine.

such as the Ordovician arandaspids. It may be a general vertebrate character, implying that the rostral region of the head developed several times. Another explanation might be that the eyeballs of the conodonts do not mark the anterior tip of the head, and that the rostral region is not preserved. In this case, the conodont apparatus, which is roughly situated beneath the eyes, may not have lain as far to the anterior as is currently believed, and the conodont elements may have been associated with the pharynx and gill apparatus, rather than with the mouth, like the pharyngeal denticles of jawed vertebrates. This would not only put conodonts closer to the jawed vertebrates, but would also explain why there is no trace of a gill apparatus more posteriorly.

Viewing conodonts as near-gnathostome animals would be consistent with the presence of cellular dermal bone or mesodentine-like tissue, shared by the gnathostomes and their closest relative, osteostracans. Moreover, the specimen described by Gabbott *et al.*⁴ seems to show a dissociation of the myomeres into dorsal and ventral parts, and this suggests the presence of a horizontal septum, yet another gnathostome character. Such an interpretation requires the assumption that conodonts have lost their exoskeleton, except in the pharynx, and their paired fins. If so, they may be regarded as derived from some jawless vertebrate that possessed pharyngeal denticles, like some of the thelodonts. The latter are an ensemble of fossil jawless vertebrates which are covered with minute scales and may include ancestors of the 'ostracoderm' groups and the gnathostomes (see figure).

Although the conodont animal is reconstructed as a small, frail, almost transparent animal that reflects the classical 'archetype' of the ancestral vertebrate, it may, after all, turn out to be a highly specialized close relative of the jawed vertebrates that developed suction feeding and pharyngeal denticles to grasp and shear prey. □

Philippe Janvier is at URA 12 du CNRS, Laboratoire de Paléontologie, 8 rue Buffon, 75005 Paris, France.

1. Briggs, D. E. G., Clarkson, E. N. K. & Aldridge, R. J. *Lethaia* **16**, 1–14 (1983).
2. Tillier, S. & Janvier, P. *La Recherche* **17**, 1574–1575 (1986).
3. Purnell, M. A. *Nature* **374**, 798–800 (1995).
4. Gabbott, S. E., Aldridge, R. J. & Theron, J. N. *Nature* **374**, 800–803 (1995).
5. Aldridge, R. J., Briggs, D. E. G., Smith, M. P., Clarkson, E. N. K. & Clark, D. L. *Phil. Trans. R. Soc., Lond.* **B340**, 405–421 (1993).
6. Aldridge, R. J. & Theron, J. N. *J. Micropaleontol.* **12**, 113–117 (1993).
7. Aldridge, R. J., Purnell, M. A., Gabbott, S. E. & Theron, J. N. *Phil. Trans. R. Soc., Lond.* **B347**, 275–291 (1995).
8. Sansom, I. J., Smith, M. P. & Smith, M. M. *Nature* **368**, 591 (1994).
9. Forey, P. & Janvier, P. *Am. Sci.* **82**, 554–565 (1994).

Bodies on the brink

Paul Weissman

It is common to think of the Solar System as ending at the orbit of the most distant planet, Pluto, at about 40 astronomical units (AU), or about 6×10^9 km from the Sun. Numerous searches for a more distant planet have all been unsuccessful. But for many years theorists have toyed with the idea that there is still something beyond Pluto, a vast belt of comets left over from the formation of the Solar System. Now, that idea is a reality.

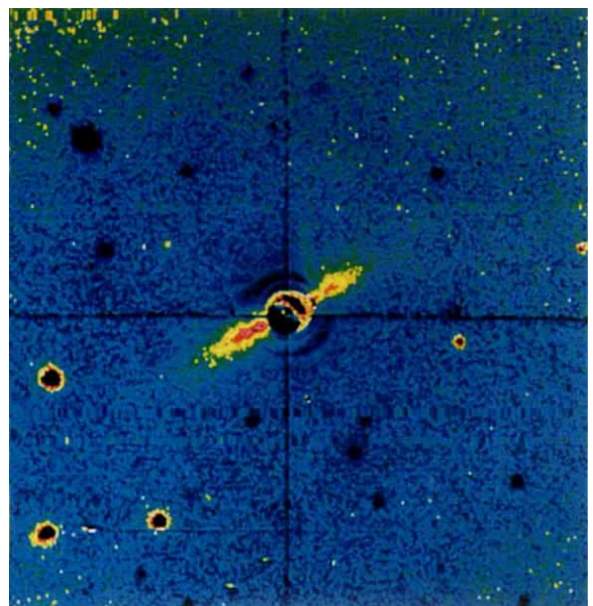
A distant comet belt beyond the planets was first suggested by K. E. Edgeworth¹ in 1949, and in a somewhat better known paper by G. P. Kuiper² in 1951. They each suggested that proto-comets formed in the solar nebula would fail to accrete into a planet at large solar distances because of the increasingly long orbital periods as one moved outwards from the Sun, and the decreasing density of nebula material. In 1980 cometary dynamicist Julio Fernandez built on this concept, suggesting that a comet belt beyond Neptune would be a dynamically efficient source for the observed short-period comets³. The short-period comets have orbital periods less than 200 years and are generally in low-inclination orbits with $i < 35^\circ$.

Subsequent computer-based simulations by dynamicist Martin Duncan and colleagues showed⁴ that the hypothetical comet belt was a far more likely source for the short-period comets than the more distant Oort cloud, the spherical cloud of some 10^{11} to 10^{12} comets surrounding the Solar System at near-interstellar distances that was proposed by astronomer Jan Oort in 1950 (see refs 5, 6). Duncan and colleagues proposed⁴ that the belt be named the 'Kuiper belt' in honour of Kuiper's 1951 paper.

The Oort cloud comets are generally between 10^4 and 10^5 AU from the Sun, and are the source of the long-period comets. According to the best current picture, the Oort cloud comets are icy planetesimals ejected from the region of the outer planets into randomly inclined, high-eccentricity orbits by the growing proto-planets. In contrast, the Kuiper belt comets are confined to low-inclination, low-eccentricity orbits beyond about 35 AU, and were never ejected because no large planet ever grew in their orbital

zone. The outer limit of the Kuiper belt is believed to be at about 10^3 AU.

The growing interest among celestial dynamicists motivated observers to intensify searches for the distant comets in the late 1980s. Among them were David Jewitt of the University of Hawaii, and Jane Luu, now at Harvard University. After five years, Jewitt and Luu's efforts were rewarded in August 1992 with the discovery of a faint, 23rd-magnitude object (about 6 million times fainter than the faintest star visible to the naked eye on a



Dusty disk seen edge-on about β Pictoris (the star itself is hidden by a circular mask).

clear night)⁷. Officially designated 1992 QB₁, the object was at 41.2 AU from the Sun, beyond Pluto's mean distance. Assuming a typical cometary albedo of 4 per cent, 1992 QB₁ is about 250 km in diameter.

An orbit solution by Brian Marsden⁸ shows that 1992 QB₁ is in a low-eccentricity orbit inclined only 2.2° to the ecliptic plane, with an orbital period of 290 years. The low eccentricity keeps 1992 QB₁ from getting any closer than about 40 AU from the Sun, and keeps it well away from the gravitational influence of Neptune. A close encounter with Neptune could throw the object into an unstable orbit that would most probably lead to its ejection from the Solar System in about 10^7 to 10^8 years. Although 1992 QB₁ could conceivably make close approaches to Pluto, that planet is too small to have much effect on its orbit. In fact, many astronomers now regard Pluto and its icy moon Charon as simply the largest objects to have accreted in the Kuiper belt.

Subsequent searches have now found a total of 23 trans-neptunian objects, 12 of them beyond 40 AU. Jewitt and Luu (along with several co-workers) have continued to be the leaders in the search for Kuiper belt objects, finding 15 of the 23, including the most distant one, 1992 ES₂ at 46.2 AU. The largest objects are each around 360 km in diameter: 1994 VK₈, found by Alan Fitzsimmons and colleagues⁹, and 1995 DC₂, found by Luu and Jewitt¹⁰.

The numbers are now sufficient to attempt some inferences on the total population of Kuiper belt comets. Jewitt and Luu do this in a paper published earlier this month¹¹. Their first seven discoveries came after searching 1.2 square degrees of sky with a CCD camera on the 2.2-metre University of Hawaii telescope on Mauna Kea. Based on these numbers, Jewitt and Luu estimate that there are at least 35,000 Kuiper belt objects with diameters greater than 100 km between 30 and 50 AU from the Sun, assuming that orbits in the Kuiper belt are confined to inclinations less than 8°. If each of these objects has a density of 1.0 g cm⁻³, then their total mass is around 2×10^{25} g, or about 0.003 Earth masses. Jewitt and Luu also noted that past searches by other astronomers limit the diameter of the largest objects within 50 AU of the Sun to about 600 km. The larger sizes of Pluto and Charon (with diameters of about 2,400 and 1,200 km, respectively) may mean that they were accretional 'run-aways', bodies that grew massive enough to begin gobbling up some of the smaller objects in the Kuiper belt.

Other estimates of the Kuiper belt population come from dynamical studies. Harold Levison and Martin Duncan, in work in progress, have used computer simulations to estimate that there must be 1.2×10^{10} comets between 30 and 50 AU, in order to provide the steady-state population of observed short-period comets in the terrestrial planets region. These comets are typically a few kilometres in diameter, much smaller than the objects currently being found by observers. Still, their total mass is estimated to be about 5×10^{26} g, or 0.08 Earth masses.

Kuiper belts are likely to be a common

feature of star formation. Between 1983 and 1984 the IRAS satellite discovered extended dust disks around several nearby main-sequence stars. In one case, the disk around the star β Pictoris was photographed edge-on, and shown to extend over 1,000 AU on either side of the central star (see photograph opposite). Shortly after the IRAS discoveries, I suggested that the source of the dust was comets in orbit around each star¹². Further analysis of IRAS data and subsequent searches have found infrared excesses indicative of disks around more than half of all main-sequence stars and also around young, forming protostars.

NEUROPSYCHOLOGY

Acting without 'seeing'

Glyn W. Humphreys

THAT 'seeing' is a complex process is evident in many ways, not least through accidents of nature which cause brain damage and produce some selective loss in the process. On page 805 of this issue¹, Castiello and colleagues report one such case, in which a patient was severely impaired at seeing more than one object — a clinical syndrome known as 'simultanagnosia'².

The particular interest of this patient, however, is that she was able to make coherent actions in response to the presence of two objects in her environment, but only when the objects were semantically related to one another — for instance, when they were two objects that might be used together or even just two objects from the same category. When the objects were unrelated, the patient's actions were incoherent, and she was unable even to align pictures of unrelated objects. Despite the effects of the related objects on her actions, the patient denied seeing both of the objects present. The results suggest that both the identity of individual objects and the relations between them in the environment can be encoded unconsciously by the visual system, so that the relations between objects can affect action even when only one of them is consciously 'seen'.

These findings add to a growing body of data in the neuropsychological literature on unconscious processing of visual information — on processing without seeing. For example, following damage to the right parietal lobe of the brain, patients may suffer the syndrome of unilateral neglect, in which they fail to respond to, and sometimes deny the presence of, stimuli that appear on the left side of space (contralateral to their lesion). Thus, when asked to decide whether there are differences between two houses, one with flames coming from a left-hand window,

The power of the repaired Hubble Space Telescope has been turned to the search for Kuiper belt objects. Search fields were taken by Anita Cochran and colleagues and are now undergoing analysis. Their preliminary results suggest that they have detected 50 to 60 objects down to diameters of 10 to 20 km, in the first field that has been examined¹³. More detailed analyses of their images are eagerly awaited. □

Paul Weissman is in the Earth and Space Sciences Division of the Jet Propulsion Laboratory, Pasadena, California 91109, USA.

such patients may respond at chance level. But when asked which house they prefer, the patients respond consistently to one house, indicating 'unconscious' processing of the differences between the houses^{3,4}. In the syndrome of visual agnosia, patients are unable to identify stimuli visually. In some cases, the damage may be so severe as to disrupt even simple visual tasks, such as matching the orientation of two lines or the sizes of patterns.

Milner, Goodale and colleagues⁵ have reported one such agnosic patient who, despite being unable to discriminate the size or orientation of stimuli in simple matching tasks, was able to reach appropriately to grasp the stimuli. The grasping reactions indicated accurate computation of the size and orientation of the patterns. In this case, conscious visual discriminations were markedly impaired, yet there was evidence that relevant visual information, computed unconsciously, was made available to the action system to direct grasping actions. Castiello and colleagues' patient also provides evidence that visual information can be used by the action system even when it is not available for conscious inspection, although in this case it seems that the information computed unconsciously includes the meaning and relations between objects.

One way of understanding these results is in terms of a distinction between separate processing routes in vision, not all of which lead to conscious object recognition and seeing. The route leading to conscious recognition and seeing probably involves 'ventral' pathways from the occipital to the temporal cortex; in addition, however, a 'dorsal' route, leading from the occipital lobe to the parietal lobe, may be involved in visually directed actions⁶. When conscious recognition of visual objects is eliminated (as in the patient of

1. Edgeworth, K. E. *Mon. Not. R. astr. Soc.* **109**, 600–609 (1949).
2. Kuiper, G. P. in *Astrophysics* (ed. Hynek, J. A.) 357–424 (McGraw-Hill, New York, 1951).
3. Fernandez, J. A. *Mon. Not. R. astr. Soc.* **192**, 481–491 (1980).
4. Duncan, M. et al. *Astrophys. J.* **328**, L69–73 (1988).
5. Oort, J. H. *Bull. astr. Inst. Netherlands* **11**, 91–110 (1950).
6. Weissman, P. R. *Nature* **344**, 825–830 (1990).
7. Jewitt, D. & Luu, J. *Nature* **362**, 730–732 (1993).
8. Marsden, B. G. *IAU Circ. No.* 5855 (1993).
9. Fitzsimmons, A. F. et al. *Minor Planets Elec. Circ. No.* 1995-C06 (1995).
10. Luu, J. & Jewitt, D. *Minor Planets Elec. Circ. No.* 1995-E07 (1995).
11. Jewitt, D. C. & Luu, J. X. *Astr. J.* **109**, 1867–1876 (1995).
12. Weissman, P. R. *Science* **224**, 987–989 (1984).
13. Cochran, A. et al. *IAU Circ. No.* 6163 (1995).

Milner *et al.*) or limited to one object at a time (as in that of Castiello *et al.*), the dorsal route may continue to operate.

The case reported by Castiello *et al.* suggests that there can be analogous limitations in the dorsal route, in making actions to more than one object at a time. Nevertheless, these limitations can be overcome if there are meaningful relations between the objects present. One possibility here is that unconscious processing of both objects present, which may still occur in the ventral route, influences the programming of actions in the dorsal route, enabling coherent actions to be made to more than one object. That is, although visual-processing routes may be separately affected by brain damage, they may normally interact so that the meaning of objects (computed by the ventral route) influences action (determined by the dorsal route). In brain-damaged patients, residual effects of such

interactions may be apparent.

Such findings indicate that seeing is not a unitary process, but that it is subserved by several separable modules. A major question for future research, and one highlighted by the new work¹, is to understand how such modules cooperate to enable coherent perception to take place. □

Glyn W. Humphreys is in the Cognitive Science Research Centre, School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.

1. Castiello, U., Scarpa, M. & Bennett, K. *Nature* **374**, 805–808 (1995).
2. Humphreys, G. W. & Price, C. J. *Cogn. Neuropsychol.* **11**, 393–434 (1994).
3. Marshall, J. C. & Halligan, P. W. *Nature* **336**, 766–767 (1988).
4. Bisiach, E. & Rusconi, M. L. *Cortex* **26**, 643–649 (1990).
5. Milner, A. D. *et al.* *Brain* **114**, 405–428 (1991).
6. Ungerleider, L. G. & Mishkin, M. in *Analysis of Visual Behavior* (MIT Press, Cambridge, MA, 1982).

MOLECULAR TECTONICS

Designer zeolites

Michael D. Ward

CRYSTAL engineering strategies for making purpose-built molecular solids hinge on constructing solid-state lattices by careful choice of the molecular structures of their components and attention to the interactions that hold the molecules in place. This approach promises unprecedented flexibility in the design of materials with desirable properties including metallic conductivity, superconductivity, nonlinear optical behaviour and ferromagnetism. A particularly elusive, yet attractive, goal has been the formation of stable molecular networks with nanometre-scale voids. As Jeffrey Moore described in a recent News and Views article¹, such networks would be molecular analogues of inorganic zeolites, which have pores of uniform size formed from rigid aluminosilicate networks. As in zeolites, voids in molecular solids could accommodate guest molecules with desirable optical properties, separate small molecules by size or chemical affinity, or provide tailored reaction environments. Unlike in zeolites, though, the forces holding molecular lattices together are generally weak, and nanoscale voids commonly collapse if emptied of reinforcing guest molecules.

As they show on page 792 of this issue, however, Gardner and co-workers² have now devised a robust molecular network with open channels which reversibly exchange solvent molecules. Their discovery illustrates that molecular solids can be designed based on topologically similar inorganic systems.

Strong and highly directional Si–O and

Al–O bonds inhibit the collapse of zeolitic networks, whereas the crystallization of molecular solids with open network structures is typically frustrated by nature's tendency to avoid empty space. Nevertheless, this has not dissuaded crystal engineers from pursuing low-density molecular networks. One approach has been to use molecules with clumsy shapes that cannot be conveniently packed. Another approach involves organic host–guest materials in which guest molecules are trapped in cylindrical, continuous channels formed by host molecules, such as urea and tri-ortho-thymotide^{3,4}.

A third strategy, which has met with reasonable success, is to use polytopic molecular building blocks which have functional groups capable of highly directional intermolecular bonding, such as hydrogen bonding or symmetric ligation of metal centres. For example, adamantane-1,3,5,7-tetracarboxylic acid crystallizes to form open diamond-like networks in which each molecule is linked to four other molecules by hydrogen bonding between the carboxylic acid groups⁵. Similarly, trimesic acid (benzene-1,3,5-tricarboxylic acid) forms a hexagonal network as a result of its threefold topology⁶. A single such network would have large voids because the volume occupied by each molecule is large relative to the connectors. But in both cases the void space is filled by topologically equivalent interpenetrating networks; the structure of adamantane-1,3,5,7-tetracarboxylic acid is actually described by five interpenetrating diamond-

like networks. Consequently, these motifs resemble inorganic materials such as potassium dihydrogen phosphate and α -tetragonal boron, $B(B_{12})_2$, which are also constructed from interpenetrating diamond-like networks.

Molecular networks have also been prepared from rigid ligating molecules and metal ions⁷, and this is the approach followed by Gardner and co-workers. The open trigonal structure that they have constructed is based on a 'hinged' network in which silver ions are coordinated to rigid, symmetric ligand molecules with threefold symmetry. The silver ions are trigonally coordinated to the cyano groups of either 1,3,5-tricyanobenzene (TCB) or 1,3,5-tris(4-ethynylbenzonitrile)benzene (TEB). Each network is an ordered array of large trigonal cavities, although these cavities contain triflate (trifluoromethylsulphonate, $CF_3SO_3^-$) counter-ions.

In the case of TCB, the trigonal layer networks stack in a staggered fashion, eliminating the continuous hexagonal channels that would result if they were in registry. In this case, the stacking motif is a consequence of the triflate ions coordinating to silver ions in the adjacent layer. Even in the absence of such an interaction, staggered stacking of layers can be a problem as adjacent layers will generally align so that the protrusions of one layer penetrate the hollows of the next⁸.

The TEB ligand also affords trigonal networks, but its larger size leads to larger trigonal voids in the network which cannot be totally filled by the triflate ions. Rather, the void space is partially filled by portions of five other topologically equivalent networks and benzene solvent molecules. The networks are held together by face-to-face π -stacking interactions between the ligand benzene rings of the different networks, preventing offset stacking. Here, instead, continuous, one-dimensional channels containing benzene solvent molecules are formed. Somewhat similar interactions have been observed in the molecular network of a manganese complex of *N,N'*-p-phenylenedimethylenebis(pyridin-4-one)⁹, in which independent interpenetrating networks are held together by π -stacking of benzene residues as well as edge-to-face (T-shaped) interactions.

The work of Gardner *et al.* is noteworthy for at least two reasons. First, they used the strategy of designing the networks around homeotypic inorganic analogues (CaCuP and LaPtSi), which have similar topological connectivity. Still more interesting is the ability of the TEB complex to exchange solvent molecules without collapsing, in this case C_6D_6 for C_6H_6 . Similar behaviour has been reported for a porous hydrogen-bonded network in which solvent molecules were incorporated into voids^{1,10}. The work of

Gardner *et al.* demonstrates that this property may be more general than previously thought, and suggests that dynamic exchange of guest molecules is feasible in porous molecular solids without network collapse.

The robustness of these porous networks during exchange is probably due to the reinforcement provided by the complex intertwining of the six independent networks, which in turn are stabilized by the closely packed π -stacks of the benzene residues. Network collapse also may be avoided if the exchange involves simultaneous exchange of molecules in the channels rather than loss followed by insertion. The large size of the channels could favour this. Further work to characterize the exchange in the nanopores may clarify the issue.

A limiting feature of molecular-based networks is their low thermal stability compared with inorganic analogues. Nevertheless, molecular-based systems may provide greater diversity in size, geometry and chemical environment of nanopores. New nanoporous materials will add to the collection of host-guest materials that have been realized with other molecular systems (such as urea and thiourea). The incorporation of properly chosen guests may lead to new materials with novel electronic properties, as happened with nanoporous zeolitic hosts.

The creation of these molecular networks, and their dynamic exchange properties, support the idea that molecular solids can be designed for specific applications, such as separations and catalysis, that require selective incorporation of molecular species into nanometre-scale voids. The topologies expected for individual networks can generally be surmised from the symmetry of the components. Predicting how these networks will interleave, though, remains difficult. As more materials of this type are produced and characterized, it is likely that the principles governing lattice assembly will become more evident. □

Michael D. Ward is in the Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, Minnesota 55455, USA.

Bananaquit biogeography

Jared M. Diamond

A STUDY by Seutin *et al.*¹, published in *Evolution*, brings modern molecular methods to the investigation of island biogeography. This of course is the field that Alfred Russel Wallace founded in 1876, after he and Darwin had recognized natural selection. The new approaches uncover character variation undetected by traditional methods, and thereby strengthen the evidence for evolutionary variation in dispersal ability.

Interests of biogeographers include infraspecific relationships among populations, from which one can potentially draw conclusions about phenomena such

mitochondrial DNA (mtDNA) among 170 bananaquits from 13 islands and five mainland sites. They come up with several interesting conclusions.

Bananaquit populations exhibit marked genetic variation. The most distinct population diverges by 2.9 per cent from other populations in its mtDNA, while several further populations diverge by 1.4 per cent. In the temperate zones, where most genetic studies of birds have been performed, such high levels of divergence characterize related species rather than subspecies, so here is more evidence that infraspecific divergence tends to be greater in tropical than temperate birds. One explanation for this could be the greater sedentariness of tropical compared with temperate birds.

Genetic variability within an island population tends to be greater on large islands than on small ones. Possible contributing factors include the greater range of habitats and altitudes (and hence of optimal genotypes) on larger islands, or the bigger population sizes and thus more rare alleles retained. Genetic variability is especially low in bananaquits of the Pearl Islands, which lie on the continental shelf off Panama and were joined to the mainland until isolated by rising sea level within the past 10,000 years. All Pearl Island bananaquits analysed shared the same haplotype. The population may have lost its variability through genetic drift without much gene input from immigrants.

Geographical variation in bananaquit plumage and mtDNA generally tally, but the new studies yield some fresh insights. For example, in its DNA the bananaquit population on Jamaica is by far the most distinctive, a conclusion that would not have been expected from plumage-based subspecific characters. The population of Puerto Rico is also somewhat distinctive, even in comparison to that on the Virgin Islands only some 70 km distant. In contrast, the bananaquits on many islands of the north-central Antilles chain are very similar, the commonest haplotype being fixed or at high frequency on all the islands. Evidently, these latter populations are the joint product of a recent range expansion and may still be connected by frequent inter-island immigration, whereas the Puerto Rican and especially the Jamaican populations are sedentary legacies of older expansions. As Seutin *et al.* point out, "These data suggest that different populations of the same species may be in different phases of colonizing activity at a specific time. . .".

In this as in other respects, the molecular findings refocus attention on one of the

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Bird in the bush — the bananaquit *Coebebe flaveola*.

as extinction and dispersal. Island populations are especially convenient units of analysis because of their defined boundaries. Typically, the object is to deduce relationships from morphological traits, such as the colour of a bird's plumage, but such an approach has three drawbacks: generally, the genetic basis of the trait is unknown; the method cannot yield genetic distance measures; and the phenotype will vary with diet and environment, as well as with genetics. Molecular characters potentially overcome all three disadvantages.

As their study site, Seutin *et al.* picked the West Indies with its dozens of ornithologically well characterized islands. As their study species, they picked a small, colourful, abundant bird called the bananaquit, which occurs on the Central and South American mainland as well as on islands. Traditionally, 41 bananaquit subspecies are recognized, largely on the basis of plumage colour. Seutin *et al.* analysed restriction-site variation of

Planet Earth Pictures

1. Moore, J. S. *Nature News and Views*, **374**, 495–496 (1995).
2. Gardner, G. B., Venkataraman, D., Moore, J. S. & Lee, S. *Nature* **374**, 792–795 (1995).
3. Otto, J. *Acta cryst.* **B28**, 543 (1972).
4. Eaton, D. F., Anderson, A. G., Tam, W. & Wang, Y. J. *Am. chem. Soc.* **109**, 1886 (1987).
5. Ermer, O. J. *Am. chem. Soc.* **110**, 3747–3754 (1988).
6. Duchamp, D. J. & Marsh, R. E. *Acta cryst.* **B25**, 5–19 (1969).
7. Hoskins, B. F. & Robson, R. J. *Am. chem. Soc.* **112**, 1546–1554 (1990).
8. Perlstein, J. J. *Am. chem. Soc.* **116**, 11420–11432 (1994).
9. Goodgame, D. M. L., Menzer, S., Smith, A. M. & Williams, D. J. *Angew. Chem. int. Edn Engl.* **34**, 574–575 (1995).
10. Wang, X., Simard, M. & Weust, J. D. J. *Am. chem. Soc.* **116**, 12119–12120 (1994).

main conclusions of classic island biogeographic studies: that there are often considerable differences in overwater dispersal ability among closely related species. For instance, the raven (*Corvus corax*) is a highly successful colonist of most British as well as Californian islands, whereas its close relatives the British and American crows (*C. corone* and *C. brachyrhynchos* respectively) rarely fly across water. Analogous differences apply even within a single superspecies or species. For instance, plumage studies of birds in the southwest Pacific show that the rufous fantail *Rhipidura rufifrons* and the golden whistler *Pachycephala pectoralis* both include highly vagile, actively expanding populations as well as sedentary, stagnant populations^{2,3}. An extreme example is Mayr's famous case⁴ of the white-eye superspecies *Zosterops griseotinctus*. The sedentary populations on the islands of Rendova and Tetipari, a mere 2 km apart, differ from each other in both plumage and voice; but another population (*Z. griseotinctus eichhorni*) is a vagile super-

tramp which is among the first colonizers of unstable small atolls and recently defaunated volcanic islands.

Dispersal ability, then, is subject to natural selection whose outcome varies according to trade-offs between costs and benefits. At least in birds, other animals and plants, dispersal ability has a genetic basis. Corresponding culturally based differences in dispersal ability distinguish island human populations. For example, the Pacific harboured both the most vagile human population in the pre-industrial world, the Polynesians, and the most sedentary of all known island human populations, the aboriginal Tasmanians. □

Jared M. Diamond is in the Department of Physiology, University of California Medical School, Los Angeles, California 90024, USA.

1. Seutin, G. *et al.* *Evolution* **48**, 1041–1061 (1995).
2. Mayr, E. & Moynihan, M. *Am. Mus. Novit.* No. 1321 (1946).
3. Galbraith, I. *Bull. Br. Mus. nat. Hist.* D4, 131–222 (1956).
4. Mayr, E. *Systematics and the Origin of Species* (Columbia Univ. Press, New York, 1942).

RIBOZYMES

Exploration by lamp light

Melissa J. Moore

If granted three wishes by the fabled genie of the lamp, proponents of the so-called 'RNA world' might ask for irrefutable evidence of three things — first, that all of the chemical reactions necessary to maintain an RNA world could have been catalysed by polyribonucleotides; second, that the present-day world of DNA, RNA and protein could have evolved from an RNA-only world; and third, that such a world did in fact exist. Now several reports, including one by Wilson and Szostak on page 777 of this issue¹, have at least partly satisfied the first two desires. The magic lamp illuminating these answers is the powerful methodology of *in vitro* RNA selection and evolution^{2,3}.

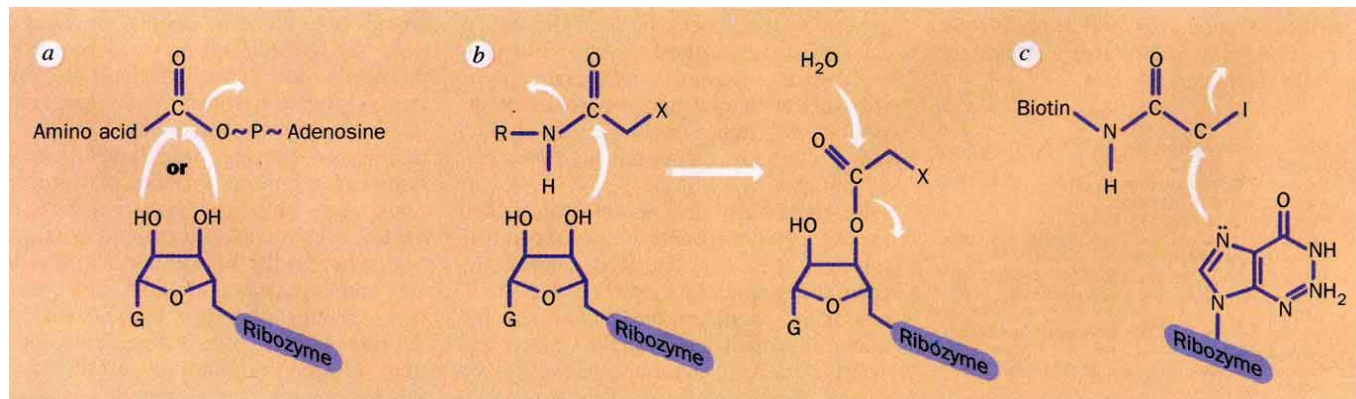
According to the RNA-world view⁴, there was a time during pre-biotic evolution when all biological reactions were catalysed by RNA. Intellectually, this model is an elegant solution to the chicken-and-egg paradox of which came first, functional proteins or information-carrying nucleic acids. In the RNA world, polyribonucleotides would have served both functions. But how many different catalytic activities would have been necessary? Certainly, RNA replicases capable of mediating template-directed polymerization of mono- or oligonucleotides would have been required for maintenance of the genetic code. But from where would the activated building

blocks used by these replicases have come? This question, of course, leaves another chicken-and-egg paradox.

On reflection, it seems reasonable that any sustainable RNA world must have had a number of biochemical pathways just for the synthesis of activated nucleosides^{5,6}. The intricacy of such pathways would have necessitated diverse catalytic activities. Although the proposal is controversial⁷, Benner *et al.*⁶ have suggested that "the last ribo-organism had a relatively complex metabolism that included oxidation and reduction reactions, aldol and Claisen condensations, trans-methylations, porphyrin biosynthesis, and an energy metabolism based on nucleoside phosphates, all catalyzed by ribo-enzymes".

What evidence is there today that RNA can catalyse such multifarious reactions? If remnants of the RNA world are to be found among contemporary ribozymes, the current fossil record is woefully incomplete. All known natural ribozymes are limited to sequence-specific cleavage or ligation of phosphodiester bonds in RNA. But the demonstration, in 1992, that a largely deproteinized ribosomal RNA retains some peptidyl-transferase activity⁸, and that the *Tetrahymena* ribozyme can slowly hydrolyse an aminoacyl ester⁹, tantalizingly hinted at the existence of a wider catalytic repertoire.

With the advent of techniques for *in vitro* selection and evolution of new RNA catalysts^{2,3}, recapitulation of the primordial ribozyme activities may at last be within grasp. The first totally *de novo* ribozymes selected, an oligonucleotide ligase¹⁰ and a polynucleotide kinase¹¹, both catalysed phosphotransfer reactions akin to those mediated by natural RNA catalysts. Last year, however, Schultz and co-workers¹² isolated an RNA that can bring about noncovalent isomerization of a bridged biphenyl ring system. And now, in 1995, we are getting the first glimpses of ribozymes that can catalyse efficient nucleophilic displacement reactions at atoms other than phosphorus.



Proposed mechanisms for displacement reactions at carbon centres catalysed by *in vitro*-selected ribozymes. *a*, Self-aminoacylation of a 3' terminal ribose. *b*, Hydrolysis of an amide bond between a deoxyribo-

nucleoside (R) and either a deoxyribonucleoside or a peptide (X). *c*, Self-alkylation of an internal guanosine. See Illangsekare *et al.*¹³, Dai *et al.*¹⁴ and Wilson and Szostak¹ for further details.

One report has described a self-aminoacylating RNA¹³. Aminoacylation is the chemical means by which a transfer RNA is charged with an activated amino acid, and RNA catalysis of this reaction would have been essential for development of even the most rudimentary translational apparatus. Yarus and colleagues¹³ selected a small RNA that can use phenylalanine-AMP as a substrate to aminoacylate its own 2'/(3') terminus (*a* in the figure). In the opposite vein, Joyce and co-workers¹⁴ reported that an evolved *Tetrahymena* ribozyme can significantly increase hydrolysis of unactivated amide bonds (*b* in the figure) like those in proteins. So not only is it possible that ribozymes could have synthesized polypeptides originally, but RNA also has the catalytic capacity to mediate a protein's demise.

Wilson and Szostak now report¹ *in vitro* evolution of the first RNA known to catalyse a bond-forming reaction on a non-nucleotide-containing substrate (*c* in the figure). Theirs is also the first example of a self-alkylating ribozyme. Analogous self-modification reactions were probably commonplace in the RNA world. By increasing the functional-group diversity available for building more and more complex active sites, RNAs containing modified sugars and bases could have achieved catalytic superiority. Thus there may have been a significant selective advantage for having modified nucleotides, as long as the modifications did not interfere with replication. Some of the 93 different modified nucleotides¹⁵ that are known in contemporary RNAs could be a legacy from this period.

Like most questions put to genies, Wilson and Szostak's selection yielded unexpected results. One is the nature of the nucleophile that the RNA has chosen. Although the alkylation reagent Wilson and Szostak used is normally specific for thiols, their self-alkylating RNA becomes modified at the N7 position of a single

internal guanosine. This is the first known chemical activation of a ring nitrogen by a ribozyme. Also striking are the dramatic structural differences between the RNA 'aptamer' first selected merely to bind the substrate, and those RNAs further evolved to perform chemistry. This demonstrates another principle when dealing with genies — obtaining an RNA that performs a certain reaction can only be guaranteed when that reaction is explicit in the selection query.

What does the future hold? Of immediate interest to enzymologists will be elucidation of the catalytic mechanisms and active-site structures of the new ribozymes. One wonders if active-site

metals will be as prominent in RNA catalysis at carbon centres as they are in reactions at phosphodiesteres. Also, to what extent will RNAs and proteins that catalyse the same reactions be mechanistically and structurally analogous? With so many research groups rubbing the magic lamp of *in vitro* selection, it is likely that RNA catalysts for many more reactions will be forthcoming shortly. The RNA-only metabolic pathways that had seemed so irretrievably lost in the past may be actually just a few years into our future. □

Melissa J. Moore is in the Graduate Department of Biochemistry, Brandeis University, Waltham, Massachusetts 02254, USA.

ASTEROIDS

243 Ida weighs in

Richard P. Binzel

FOR nearly all of the three centuries since their discovery, asteroids have been seen by telescope only as unresolved points of light floating in space. Now, all that is changing. Two reports on pages 783 and 785 of this issue^{1,2} illustrate the wealth of information to be obtained from spacecraft images and spectra. Their subject is asteroid 243 Ida, the only asteroid known to have its own moon.

Although advances in astronomical instrumentation and dynamical modelling have allowed asteroid science to mature greatly over the latter half of this century, spacecraft images have been needed to confirm the basic tenets predicted by decades of previous study: asteroids are pock-marked planetesimals remaining from the time of the Solar System's formation, which have subsequently undergone a complex history of thermal processing, collisions and orbital evolution. Spacecraft measurements of surface spectral properties, masses, densities and elemental abundances have also been needed to establish links between asteroids and the wealth of laboratory data derived from their meteorite progeny found on the Earth.

The Galileo spacecraft's flybys fulfilled optimistic expectations for mapping surface features and spectral properties of the asteroids 951 Gaspra (in October 1991) and 243 Ida (in August 1993), but the opportunity for Galileo to yield the mass and density of an asteroid came as a surprise. As Belton *et al.*² explain on page 785 of this issue, the perpetrator of the surprise is a tiny satellite, now named Dactyl, discovered in orbit around Ida in some of the images relayed last year³. Their density determination for Ida, coupled with a spectral comparison between Ida and Dactyl reported by Chapman *et al.*¹ on page 783 of this issue, is tipping the

balance in a seesaw debate on how to link the most common class of meteorites, the ordinary chondrites, with what is known of the properties of asteroids.

Ground-based studies of Ida's spectral properties place it within a broad category known as the 'S-class'. Asteroids in this class are inferred to have compositions ranging from olivine-rich to pyroxene-rich mineralogies with varying metal abundances⁴. Ida's particular S-class spectrum can be interpreted as being consistent with two different types of meteorites: ordinary chondrites and stony-irons. Although both meteorite types contain olivine, pyroxene and metal constituents, the ordinary chondrite material exists in a primitive state — it has not been substantially heated since it condensed from the solar nebula. Stony-irons, however, are distinguished by a history of sufficient heating for the minerals and metal to segregate before recrystallizing. As a consequence of the metal segregation, stony-irons have a greater density (typically about 5.1 g cm⁻³) than ordinary chondrites (about 3.6 g cm⁻³).

Measuring an asteroid's density has long been a goal for asteroid mission planners. But Galileo is not an asteroid mission. It is a Jupiter mission (with a tortuous history) whose top priority is safe arrival and scientific operations there. Determining Ida's mass directly from the spacecraft's slightly deflected flyby trajectory would have required an uncomfortably close passage. Reports of stellar occultations by asteroids, observed mostly by amateur astronomers during the past two decades, have suggested that some asteroids have satellites or debris clouds in their vicinity⁵. Although these reports remained largely unconfirmed and none concerned Ida, mission planners wanted

1. Wilson, C. & Szostak, J. W. *Nature* **374**, 777–782 (1995).
2. Joyce, G. F. *Curr. Opin. struct. Biol.* **4**, 331–336 (1994).
3. Chapman, K. B. & Szostak, J. W. *Curr. Opin. struct. Biol.* **4**, 618–622 (1994).
4. Gilbert, W. *Nature* **319**, 618 (1986).
5. Joyce, G. F. & Orgel, L. E. in *The RNA World* (eds Gestland, R. F. & Atkins, J. F.) 1–25 (Cold Spring Harbor Laboratory Press, New York, 1993).
6. Benner, S. A. *et al.* in *The RNA World* (eds Gestland, R. F. & Atkins, J. F.) 27–70 (Cold Spring Harbor Laboratory Press, New York, 1993).
7. Maizels, N. & Weiner, A. M. *Nature* **330**, 616 (1987).
8. Noller, H. F., Hoffarth, V. & Zimniak, L. *Science* **256**, 1416–1419 (1992).
9. Piccirilli, J. A. *et al.* *Science* **256**, 1420–1424 (1992).
10. Bartel, D. P. & Szostak, J. W. *Science* **261**, 1411–1418 (1993).
11. Lorsch, J. R. & Szostak, J. W. *Nature* **371**, 31–36 (1994).
12. Prudent, J. R., Uno, T. & Schults, P. G. *Science* **264**, 1924–1927 (1994).
13. Illangsekere, M., Sanchez, G., Nickles, T. & Yarus, M. *Science* **267**, 643–647 (1995).
14. Dai, X., Mesmaeker, A. D. & Joyce, G. F. *Science* **267**, 237–240 (1995).
15. Limbach, P. A., Crain, P. F. & McCloskey, J. A. *Nucleic Acids Res.* **22**, 2183–2196 (1994).

to play it safe and keep the spacecraft about 100 diameters away, a theoretical limit for temporarily bound satellites and debris. From such a distance, a direct mass determination is impossible.

The discovery of an object 1.4 km in diameter close to Ida showed how wise the flight engineers were to be cautious and served as a testimonial to the contributions made by amateur astronomers. Belton *et al.* point out that the likelihood of the object's being a random passing asteroid is astronomically small. The real asteroid belt does not resemble the crashing boulder fields depicted in popular science fiction films; on average, kilometre-sized asteroids are separated by millions of kilometres. From its slow relative motion with respect to Ida, the Galileo scientists further argued, and the International Astronomical Union agreed, that the object must be a satellite in orbit around Ida. The name Dactyl is derived from beings in Greek mythology who resided on Mount Ida.

Whereas the discovery of Dactyl might have been a gift from the Fates, determining its orbit was a task for the Titans. Although the satellite was identified in 47 separate images, Galileo, Ida and Dactyl represented three moving targets experiencing a constantly changing geometry. In addition, the duration of the encounter covered only a fraction of any likely orbital period. The task accomplished by Belton *et al.* was to constrain the orbit as well as possible, and, with the help of Kepler's Third Law, to derive a mass for Ida. Combining this mass with a volume derived from imaging, they find a bulk density for Ida of $2.6 \pm 0.5 \text{ g cm}^{-3}$, lower than expected for either pure stony-iron or ordinary chondrite material. To account for this density, Belton *et al.* argue that Ida has some internal porosity, assuming its interior is fractured. A porosity between 23 and 48 per cent is most consistent with values inferred for other small Solar System bodies. Porosities in this range yield Ida's observed bulk density if the predominant solid material is chondritic.

If Ida is indeed composed of ordinary chondrite material, why don't its spectral characteristics (or those of S-asteroids in general) more closely match laboratory spectra of chondrite meteorites? Central to the long debate⁶ on whether the S-asteroids have an ordinary chondrite or stony-iron connection has been whether some surface alteration, termed 'space weathering', might be operating on asteroids as it does on the Moon⁷. Demonstrating that such a process occurs on asteroids is tantamount to closing the link between S-asteroids and ordinary chondrites.

In comparing Ida's spectrum with its satellite's, Chapman *et al.* argue that the less red colour of Dactyl and its deeper absorption bands are evidence of space

weathering. Specifically, they assume that Ida and Dactyl originated as identical twins in the break-up of a single parent body that formed a large cluster of asteroids, known as the Koronis family, in which they reside. Dactyl being smaller, they suggest that its surface is more frequently refreshed and younger than Ida's. Ida's older and redder surface, they argue, demonstrates the occurrence of a space weathering process that could forge the sought-after link between S-asteroids and ordinary chondrites.

Although Ida's density measurement and Dactyl's colour difference tip the scales towards an ordinary chondrite link for Ida and similar S-asteroids, a final decision would be premature. The unknown porosity of Ida, which could be substantially higher than assumed if it is composed of multiple pieces reassembled from the Koronis parent body, still leaves room for argument for a stony-iron composition. Similarly, no space weathering process need be invoked if Ida and Dactyl are fraternal siblings. Dactyl's spectral signature does fall within the range seen for other Koronis family members. Even Dactyl's 'fresh' surface does not display a spectrum fully matching that of an ordinary chondrite meteorite.

Where do we go from here? The new era of asteroid science ushered in by Galileo's flybys was made possible though NASA's farsighted policy of allowing outer planet missions to investigate asteroids encountered *en route*. This policy has now been abandoned, and sadly, the Cassini mission to Saturn will turn a blind eye when it passes through the asteroid belt.

The missing ingredient in Galileo's opportunistic asteroid missions has been an elemental abundance analysis which can only be obtained by a rendezvous. Fortunately, such measurements are a goal for the Near-Earth Asteroid Rendezvous (NEAR) mission scheduled for launch next year. As part of NASA's new Discovery programme, NEAR will rendezvous with the S-type asteroid 433 Eros in 1999. From this first dedicated asteroid mission, we can hope that a knock-out punch will be delivered in resolving a meteorite analogue for the S-asteroids. □

Richard P. Binzel is in the Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

1. Chapman, C. R. *et al.* *Nature* **374**, 783–785 (1995).
2. Belton, M. J. S. *et al.* *Nature* **374**, 785–788 (1995).
3. Belton, M. J. S. & Carlson, R. *IAU Circ.* No. 5948 (1994).
4. Gaffey, M. J. *et al.* *Icarus* **106**, 573–602 (1993).
5. Van Flandern, T. C., Tedesco, E. F. & Binzel, R. P. in *Asteroids* (ed. Gehrels, T.) 443–465 (Univ. Arizona Press, Tucson, 1979).
6. Wetherill, G. W. & Chapman, C. R. in *Meteorites and the Early Solar System* (eds Kerridge, J. F. & Matthews, M. S.) 35–67 (Univ. Arizona Press, Tucson, 1988).
7. Pieters, C. M. *et al.* *J. geophys. Res.* **98**, 20817–20824 (1993).

Money teaches

THE recent British 'Science Week' aimed to persuade children to take up science, on the absurd pretence that Science is Fun. In fact there is no demand for scientists, as shown by their low salaries and dismal career prospects. Daedalus is seeking a sounder way to influence the educational choices of the young. He notes how the stockmarket enables a technical project to be judged at every stage, not by its starry-eyed proponents, but by those shrewd outsiders, the shareholders. He therefore wants to take student loan schemes to their logical conclusion, and set up a formal stockmarket in students.

A potential student will simply float himself on the educational stockmarket, and issue share capital in the project of his own education. At once the benevolent disciplines of the market will come into play. Suppose he wants to study physics. If the market feels him to be a rotten physicist, or reckons there are too many physicists already, he will find it hard to raise capital. His shareholders will steer him towards classics, advertising studies, or wherever they see the best future returns. They may even judge that he already has as much education as he can hold, and should get a job instead.

This elegant system does away with rigid formal entrance requirements. Dedicated but unqualified mature students, outsiders with unusual qualifications, eccentric geniuses who terrify the examiners, all will have a fair chance at last. They need only convince the market of their worth, and the portals of education will open to them. Once inside, they will have a powerful incentive to do well. If they seem out of their depth, or neglectful of their studies, further investment may not be forthcoming.

Colleges and universities will benefit too. At last they will have clear market signals about the worth of their activities. If the share price of the students plummets when Professor X is appointed, or the new ecosociology course is unveiled, a wise vice-chancellor will reconsider his plans. Yet many tricky decisions will be avoided. Projects to encourage minorities or women to go in for this or that subject, for example, will become irrelevant. If it's worth doing, the market will do it anyway.

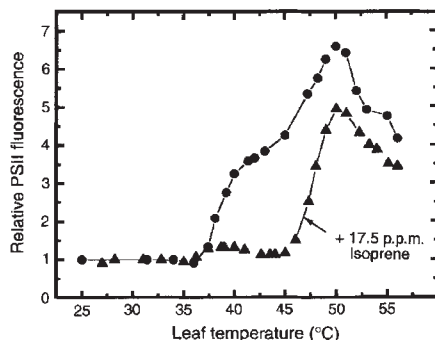
Once a student has graduated, the shareholders will look for their return. He will pay them a steady dividend on his earnings, deductible against tax (after all, the State has not paid for his education). If he does particularly well, he may even be able to buy back his own shares, and enjoy the pleasure of full self-made independence. David Jones

Why plants emit isoprene

SIR — Global emission of isoprene (2-methyl-1, 3-butadiene) from vegetation is estimated at 3×10^{14} g yr⁻¹, similar to methane¹. Because isoprene reacts very rapidly with hydroxyl radicals, it plays an important role in atmospheric chemistry²⁻⁴. Why plants emit isoprene has been a mystery since Sanadze⁵ first discovered the phenomenon in the 1950s.

Isoprene is emitted from many but not all plants. It is dependent upon photosynthesis and typically accounts for 2% of photosynthesized carbon at 30 °C in aspen and oak trees^{6,7}. Darkness, inhibitors of photosynthesis, or an atmosphere of pure N₂ will stop isoprene emission. Isoprene emission is much more sensitive to temperature than is photosynthesis. Over short periods of time, isoprene emission can increase as much as tenfold with a 10 °C increase in leaf temperature⁸. Plants grown at 20 °C or less do not emit isoprene until induced by treatment at 30 °C for 4 to 6 h. Water stress, which is often accompanied by high temperature stress, also induces plants to make isoprene⁸. These observations led us to consider whether isoprene synthesis helps plants cope with high temperature.

Isoprene provided thermal protection



Effect of exogenous isoprene on chlorophyll fluorescence from a kudzu (*Pueraria lobata* (Willd.) Ohwi.) leaf grown at 28 °C. Chlorophyll fluorescence yield was measured with a pulse-modulated fluorometer (PAM, Heinz Walz; Eppeltrich). White light was provided by a xenon-arc lamp and was $1,000 \mu\text{mol m}^{-2} \text{s}^{-1}$. Endogenous isoprene production was suppressed by conducting the experiment in N₂. Isoprene was added to the airstream by passing it through a short piece of polyvinyl chloride tubing inside an Erlenmeyer flask containing a high concentration of isoprene. Isoprene diffused through the tubing into the air stream at a constant rate. Isoprene in the air stream was measured by injecting a 10-ml sample of air into a cryogenic trap held in liquid N₂ which was then warmed to release the isoprene into a Shimadzu gas chromatograph fitted with a photoionization detector. The gas chromatograph was calibrated by mixing liquid isoprene in two stages down to 256 p.p.b. in N₂. Thermal protection has been seen in measurements made with more than 20 pairs of leaves under various conditions, including darkness.

in a range of conditions, including darkness. We found the greatest degree of thermal protection when leaves were exposed to $1,000 \mu\text{mol m}^{-2} \text{s}^{-1}$ photons (half of sunlight) and a nitrogen airstream (to suppress isoprene emission). The increase in chlorophyll fluorescence which indicates irreversible leaf damage⁹ occurred at 37.5 °C in a leaf without isoprene but it occurred at 45 °C in a leaf given 17.5 p.p.m. isoprene in the airstream (see figure). In normal air, we estimate that the concentration of isoprene inside leaves at 40 °C is typically between 1 and 20 p.p.m. and is much higher during water stress. We have demonstrated thermal protection in experiments where photosynthesis was measured as CO₂ uptake and we have been able to demonstrate dose dependence. We speculate that isoprene could dissolve into membranes and alter their properties or alter membrane-protein interactions to increase thermal tolerance. Other isoprenoid compounds are also important in membrane physiology¹⁰.

Because isoprene is made rapidly and also lost rapidly from the leaf, its concentration will be well correlated with leaf temperature through the day. The high rate of synthesis, coupled with the volatility, provides a mechanism for rapidly increasing the isoprene concentration inside leaves when they heat up and decreasing the concentration upon leaf cooling. This could be advantageous in environments with fluctuating tempera-

ture. It also means that the greatest rates of isoprene emission from vegetation are likely to occur on hot, still days when there is a potential for high levels of tropospheric ozone in polluted air. Ozone abatement has been based primarily on reduction of anthropogenic hydrocarbons, but given the large amount of biogenic isoprene released on hot days, control of anthropogenic NO_x may be more effective.

Thermal tolerance resulting from isoprene can be substantial and the concentration of isoprene required is within the range expected in plants. The temperature-sensitivity of isoprene synthesis and the inducibility by high temperature and water stress support our hypothesis that thermal tolerance explains why many plants emit isoprene. We believe that thermal tolerance may be the primary explanation for why plants emit isoprene.

Thomas D. Sharkey

Eric L. Singaas

Department of Botany,

University of Wisconsin-Madison,

Madison, Wisconsin 53706-1383, USA

1. Brasseur, G. P. & Chatfield, R. B. in *Trace Gas Emission from Plants* (eds Sharkey, T. D., Holland, E. A. & Mooney, H. A.) 1-27 (Academic, San Diego, 1991).
2. Trainer, M. et al. *Nature* **329**, 705-707 (1987).
3. Chameides, W. L., Lindsay, R. W., Richardson, J. & Kiang, C. S. *Science* **241**, 1473-1475 (1988).
4. Thompson, A. M. *Science* **256**, 1157-1165 (1992).
5. Sanadze, G. A. in *Trace Gas Emissions by Plants* (eds Sharkey, T. D., Holland, E. A. & Mooney, H. A.) 135-152 (Academic, San Diego, 1991).
6. Monson, R. K. & Fall, R. *Pl. Physiol.* **90**, 267-274 (1989).
7. Loreto, F. & Sharkey, T. D. *Planta* **182**, 523-531 (1990).
8. Sharkey, T. D. & Loreto, F. *Oecologia* **95**, 328-333 (1993).
9. Seemann, J. R., Berry, J. A. & Downton, W. J. S. *Pl. Physiol.* **75**, 364-368 (1984).
10. Ourisson, G. & Nakatani, Y. *Chem. Biol.* **1**, 11-23 (1994).

Exotic pests and parasites

SIR — Orr *et al.* in Scientific Correspondence¹ describe the competitive dominance of the red imported fire ant *Solenopsis invicta* at food resources in Brazil, which is diminished when under attack by phorid flies, and suggest that this parasite could be used to control the ant in the southern United States. We can add an interesting note to the American story of *S. invicta*, and comment on the role of parasitism in structuring natural communities.

S. invicta might have escaped regulation by its coevolved parasitic organism in an alien ecosystem when it was introduced into the southern United States in the 1940s, but at some time after this it acquired another exotic parasite species, the strepsipteran *Caenocholax fenyesi*², first described in Brazil in 1904. This parasite is common in South America, but its host in endemic areas is still unknown. We speculate that the strepsipteran was introduced into the southern United States with another host ant species, but now commonly occurs in the red imported fire

ant. This is a good example of an introduced species (the red imported fire ant) which has become aggressive and abundant in an alien ecosystem where its natural parasites (including the phorid fly) are absent. Now the niche left by the natural parasite seems to have been taken over by a new parasite (the strepsipteran). It is to be hoped that this parasite may be followed by others (for example, microbial, introduced phorids) and a balance will begin to be restored.

C. fenyesi belongs to the curious strepsipteran family Myrmecolacidae, the males of which parasitize hosts belonging to a different order from the hosts of females. In the New World the male and female and their hosts of only one species are known³; of the 87 species of Myrmecolacidae described worldwide, five have been females and only two of these have been associated with their males. Myrmecolacidae are more numerous and speciose in the New World than the Old. The only species of Myrmecolacidae found in the southern United

States is *C. fenyesei*, and its range is confined within that of its seemingly new host, *S. invicta*. Based on the five females described so far, we believe that the host of female *C. fenyesei* will turn out to be an orthopteran. But the female of *C. fenyesei* in the southern United States is likely to have switched hosts, just as the male has done.

A safer course than the suggestion by Orr *et al.* of the introduction of *S. invicta*-specific phorid flies to the southern United States (which has the risk that these too might switch hosts and damage other species in an alien ecosystem) would be to hunt for the females of *C. fenyesei* and use them for the dispersal of the parasites to control the red imported fire ant. Adult hosts parasitized by Strepsiptera are unable to reproduce and are killed only after the emergence of the strepsipteran⁴; hence while technically a parasite (by host survival), a strepsipteran is effectively a 'parasitoid'⁵. Males and young queens are among the individuals

affected. But as it most commonly affects the major workers, which are sterile anyway, *C. fenyesei* is perhaps best considered as a macroparasite of a superorganism, the ant colony, which it debilitates in all the above mentioned ways.

In a curious inversion on the other side of the globe, a hunt is now on for the males of another Myrmecolacidae, *Stichotrema dallatorreanum* (presumably in an ant), the females of which parasitize and significantly reduce the long-horned grasshopper *Sexava* sp., a severe pest of oil palm in Papua New Guinea.

J. Kathirithamby

W. D. Hamilton

Department of Zoology,
University of Oxford,
Oxford OX1 3PS, UK

1. Orr, M. R. *et al.* *Nature* **373**, 292–293 (1995).
2. Kathirithamby, J. & Johnston, J. S. *Ann. ent. Soc. Am.* **85**, 293–297 (1992).
3. Ogloblin, A. A. *Proc. Int. Cong. Ent.* **2**, 1277–1284 (1939).
4. Kathirithamby, J. *Syst. Ent.* **14**, 41–92 (1989).
5. Kathirithamby, J. & Hamilton, W. D. *Trends ecol. Evol.* **7**, 349–351 (1992).

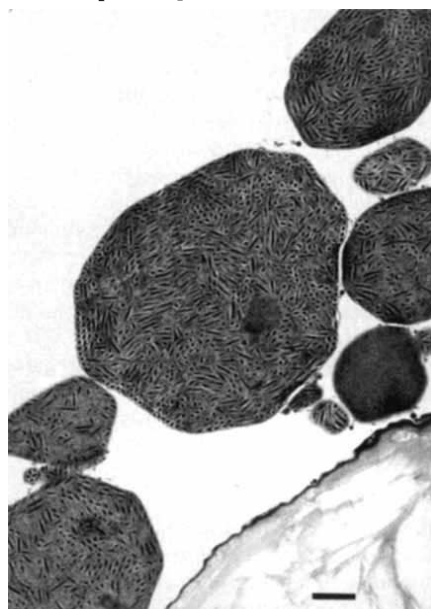
Gonad-specific virus of corn earworm

SIR — Two closely related species of moths, the corn earworm *Helicoverpa zea*, in the United States, and the legume pod borer/cotton bollworm *H. armigera*, in Asia, Africa and parts of Europe, constitute major pests of several field crops. In a colony of *H. zea*, maintained at Stoneville, Mississippi, we observed atrophy of the reproductive system in 20–55% of adults (mean 35% over seven generations), with no mating observed among such adults.

Here we report the discovery of a previously undescribed virus associated with these agonadal (AG) moths. The AG females have grossly deformed common and lateral oviducts which are almost 100 times the size found in normal females. The common oviduct is full of a white buttery mass of virus-atypical occlusion bodies. The females also have no ovaries, bursa copulatrix, accessory glands or spermatheca. The AG-males have a normal-looking endophallus, aedeagus and ductus ejaculatorius simplex. The testes do not fuse as in normal males and are very small (about the size found in the third-instar larvae) whereas seminal vesicles, vasa deferentia, duplexes and accessory glands are absent. Atypical occlusion bodies are mainly present in the posterior portion of the ductus ejaculatorius simplex. In most AG-adults, the rectum is greatly swollen and filled with fluid containing many bacteria.

Ultrastructural studies have shown that the virus is confined to the reproductive system; therefore we designate it a gonad-specific virus (GSV). In the female, the lumen of oviduct was full of virus-atypical occlusion bodies, with no visible cytopathology of the adjoining cells. The

atypical occlusion bodies differ from those of the typical nuclear polyhedrosis virus (NPV) in that they contain high concentrations of virions and have a granular matrix rather than a typical polyhedrin protein matrix. The occlusion bodies also appear to have a host-derived 'membrane' rather than a virus-derived polyhedron-membrane. No polyhedra typical of *H. zea* NPV (ref. 1) were observed in various tissues (tracheal matrix, hypodermis, fat bodies, oviduct and Malpighian tubules) of GSV-infected insects. The virus was found replicating in nuclei of the cells in



Atypical occlusion bodies of the gonad-specific virus in the lumen of the oviduct of the corn earworm *H. zea*. Each occlusion body contains many virions. Scale bar, 1 μ m.

the distal part of the oviduct. Virions were also seen in the cytoplasm from which clumps of these virions budded into the lumen of the oviduct. The nucleocapsids measured $382 \pm 30 \times 77 \pm 3$ nm ($n = 50$). By contrast, nucleocapsids produced in the typical *H. zea* NPV infection in fat body tissues² measured $318 \pm 18 \times 42 \pm 4$ nm ($n = 50$). Preliminary biochemical tests using oligo primers designed to detect various baculovirus genes including the highly conserved polyhedrin gene from the NPV of gypsy moth indicate that GSV is different (E. Dougherty, personal communication). If GSV is a baculovirus, it represents a unique phenotype.

Unlike other known insect baculoviruses, GSV caused no larval mortality, and low (average 26%) AG condition in adults that were fed as first-instar larvae on a diet smeared with the virus. When newly emerged females were fed on a 10% sucrose solution containing the crude virus and then allowed to mate on the following day, 25% females and 40% males in the resulting progeny were AG. Injection of a crude preparation of GSV at doses as low as 0.001 oviduct equivalents into the abdomen of newly emerged females, and mating of such females to normal males on the following day, resulted in progeny that was >95% AG. Because immersing *H. zea* eggs in a crude virus solution did not cause the AG condition in the resulting adults, the virus apparently penetrates the eggs before their chorion is hardened prior to oviposition. The virions of GSV could also be carried by the sperm into the egg at the time of fertilization. These observations suggest a transovarial mode of transmission for GSV.

Normal-looking females among the GSV-infected progeny do carry the virus, perhaps in a low titre, as shown both by electron microscopy of their oviducts as well as by injecting the oviduct homogenates into newly emerged females and examining their progeny. It is therefore feasible that GSV could be introduced into the natural population of *H. zea* by infecting newly emerged females with a very low titre of the virus and releasing these females into the field. Spraying a viral suspension on the plants or providing virus-laced sucrose feeding stations in the fields may be additional methods for dispersing the virus among natural populations. But these possibilities must first be evaluated in field tests.

Ashok K. Raina

Insect Neurobiology and
Hormone Laboratory,

Jean R. Adams

Insect Biocontrol Laboratory,
USDA Agricultural Research Service,
BARC-East, Beltsville,
Maryland 20705, USA

1. Ignoffo, C. M. *Expl. Parasit.* **33** 380–406 (1973).
2. Adams, J. R. & McClintock, J. T. in *Atlas of Invertebrate Viruses* (eds Adams, J. R. & Bonami, J. R.) 87–204 (CRC, Boca Raton, Florida, 1991).

Electrically conductive Langmuir–Blodgett films of charge-transfer materials

Martin R. Bryce & Michael C. Petty

The Langmuir–Blodgett technique allows the controlled layer-by-layer deposition of ordered molecular films. When such films are fabricated from charge-transfer compounds, they can be electrically conducting, and may find applications in thin-film molecular electronic devices.

THERE is considerable scientific and technological interest in the design, synthesis and characterization of organic materials that possess unusual solid-state properties. Of particular interest are organic charge-transfer (CT) salts^{1–8}, a class of materials in which metallic conductivity—and even superconductivity—has been observed. The motivation for this research is essentially twofold. On a fundamental level, there is a wealth of new solid-state chemistry and physics to be uncovered. More practically, there is the potential for technological applications in the rapidly expanding field of molecular electronics.

Studies of conducting charge-transfer salts have concentrated almost exclusively on single crystals. But single crystals, although providing 'clean' systems for investigating the underlying physical processes, are both fragile and extremely difficult to process; this greatly limits the prospects for practical applications. Work on charge-transfer salts is now moving in an exciting new direction, with the preparation of these materials in the form of thin Langmuir–Blodgett (LB) films. The LB technique (Box 1) provides a means by which organic molecules can be assembled with a greater level of structural control than in the corresponding crystalline materials, and conducting CT salts in the form of thin films should be much easier to incorporate into practical device structures.

Here we review the progress that has been made to date, highlighting those developments that appear especially promising, and discussing some of the issues that will need to be addressed before the technological potential of these materials can be realized.

Conductivity and superconductivity in CT salts

The electrical properties of CT salts in the form of LB films differ in many respects from those of their single-crystal counterparts, and remain poorly understood. Nevertheless, the general picture that has emerged from work on single crystals provides a useful framework for interpreting the results of thin-film studies.

Crystals of conducting CT salts comprise highly ordered arrays of donor and acceptor species, one or both of which must be a radical ion that is thermodynamically stable. Orbital overlap provides a mechanism for delocalization of electrons, with the width of the resulting electronic bands dependent on the degree of interaction between the orbitals on neighbouring molecules. But extensive interaction of molecular orbitals is not in itself sufficient to produce metallic (or even semiconducting) behaviour: the size of the energy gap between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) is critical (Fig. 1). (The HOMO and LUMO form, respectively, the valence and conduction bands of the material.) When this gap is large, the material is an insulator. As the band gap decreases, thermal excitation of electrons from the valence band to the conduction band becomes more effective and the material becomes an intrinsic semiconductor. True metallic behaviour, on the other hand, requires partially filled bands, in which it is possible for a large number of electrons to move into infinitesimally higher energy states. In this case, it is the electrons

in or near to the highest occupied energy level (the Fermi level) that determine the electrical properties. In CT salts, the band states are derived from the HOMO of the donor species and the LUMO of the acceptor species. The conductivity of organic (and inorganic) metals, in the absence of impurities and defects, is limited by the scattering of electrons by vibrations of the atomic lattice (phonons); as the temperature is lowered there are fewer lattice vibrations and the conductivity increases. In contrast, the temperature dependence of the conductivity of a semiconductor is dominated by the energy gap; as the temperature is lowered, there is less energy available to excite charge carriers from the valence band to the conduction band, and conductivity decreases.

Superconductivity in inorganic materials stems from the highly correlated motion of electron pairs (Cooper pairs). This mechanism is clearly distinct from metallic conduction for which individual electrons are the charge carriers and cannot be predicted from the simple band scheme outlined above. Electron pairing in conventional superconductors is mediated by phonons below a critical temperature, T_c , only if a strict set of structural and electronic energy conditions is fulfilled; however the criteria for superconductivity in organic materials are not yet known.

To obtain efficient intermolecular interactions in a CT salt, the structural and redox properties of the constituent molecules are clearly important³. Some of the most commonly used donor and acceptor species are shown in Fig. 2. The donor and/or acceptor molecules are typically planar (or nearly planar), and form either segregated stacks (Fig. 3) or planar sheets. The electrical properties of the resulting salt are highly anisotropic. The electronic properties of such low-dimensional systems were first discussed over 30 years ago. Frohlich⁹ and Peierls¹⁰ pointed out that, at low temperatures, a quasi-one-dimensional metal, in which the atoms are uniformly spaced, would be unstable with respect to lattice distortions (the chemical analogue is the well-known Jahn–Teller distortion). The effect of this distortion is to lower the symmetry of the system, which in turn influences the electronic band structure. For CT salts, the degree of instability depends on the level of band filling (effectively the number of radical ionic and neutral molecules in the stack). A half-filled band has no neutral molecules in the stack, and as each molecule has an unpaired spin, there is an electronic driving force for spin pairing. In consequence, the molecules in the stack can dimerize, concomitant with the creation of an energy gap between bonding and antibonding energy levels, and metallic conduction is lost (Fig. 4A, B). This is the well known Peierls distortion, and leads to alternating regions within the lattice of higher and lower charge density—a charge density wave (CDW). The introduction of neutral molecules into the conducting stack reduces the level of band filling and the CDW is therefore free to move along the stack and thus act as a charge carrier. (Although qualitatively correct, this model is an oversimplification. In the case of the CT salt TTF–TCNQ, the donor and acceptor molecules form separate conducting stacks with very different electronic band widths. Although the periodicities of the CDWs are incommensurate with the lattice on the individual stacks,

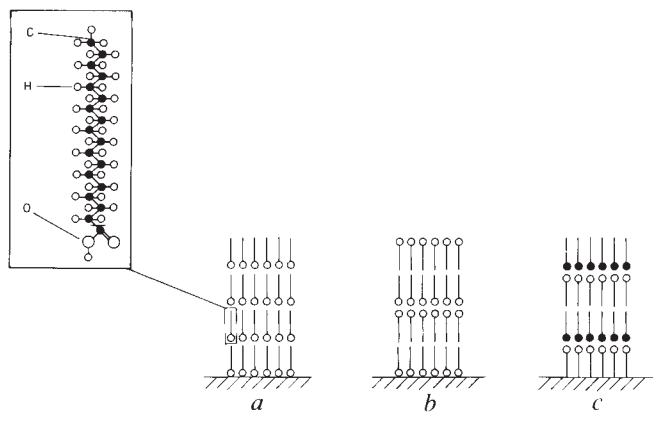
BOX 1 The Langmuir-Blodgett technique

LANGMUIR-BLODGETT films are prepared by first depositing a small quantity of an amphiphilic compound (that is, one incorporating both polar and nonpolar groups—the 'classical' materials being long-chain fatty acids), dissolved in a volatile solvent, onto the surface of purified water. When the solvent has evaporated the organic molecules may be carefully compressed to form a floating 'two-dimensional' solid. In fact, monolayer films can exhibit a multiplicity of phases as they are compressed on the water surface; these are similar to the various mesophases shown by liquid-crystal materials. If the surface pressure (surface tension) of the monolayer is held constant in one of the more condensed phases, then the film may be transferred from the water surface onto a suitable solid substrate simply by raising and lowering the latter through the monolayer-air interface (see figure, which shows the molecular arrangements of amphiphilic carboxylic acid molecules). A number of deposition modes are possible depending on the interactions between the polar and nonpolar parts of the amphiphiles, and the nature of the bond between the first layer and the substrate surface. These are: X-type deposition (film transfer on the downstroke only); Z-type (transfer on the upstroke only)—a in the figure; and Y-type (transfer on both the upstroke and downstroke)—b in the figure. Film deposition may usefully be characterized by the decrease in the area occupied by the monolayer on the water surface divided by the coated area of the solid substrate (the deposition, or transfer, ratio).

It is also possible to construct LB films using more than one type of molecular film. In the simplest case, condensed floating monolayers of two different amphiphilic materials are confined (using mechanical barriers) to different regions of the water surface. By lowering the solid substrate through the first layer, of, say, material A, and raising it up

through the other, material B, alternate-layers of structure ABABA... may be built up—c in the Figure.

The control of the molecular architecture is intrinsic to the LB process and is clearly one of the most attractive features when applied to the formation of multilayers (which, in effect, are segregated stacks) of π -donor and π -acceptor species. Moreover, a significant advantage of alternate-layer films, over solution-cast films or crystallized samples, is that the molecules can, in principle, be aligned in a non-centrosymmetric array and can, therefore, exhibit physical phenomena such as pyroelectricity and second-harmonic generation.



they can nevertheless order with respect to each other. This 'interstack' ordering provides an alternative mechanism for the Peierls instability, and leads to the formation of an insulating state below ~ 40 K (refs 2, 5).

In low-dimensional conductors, metal-insulator transitions can also be driven by spin density waves (SDW) or anion ordering². The former instability arises when the electron spins order antiferromagnetically (alternating spin 'up' and spin 'down', Fig. 4C) which limits conductivity by introducing an energy gap in the electronic band. This effect is seen in crystals of TMTSF salts. Application of external pressure may suppress this distortion^{4,5} and, for some systems, can result in a superconducting ground state. But why a magnetic (SDW) transition and not a Peierls (CDW) transition should occur in these systems is not clear at present. The effect of anion ordering on metal-insulator transitions was first recognized in the CT salts (TMTSF)₂⁺ X⁻ (where X is PF₆⁻, AsF₆⁻, ReO₄⁻ or ClO₄⁻) (ref. 5). At room temperature the anions have random orientations, whereas at low temperature they orient to give three-dimensional superstructures (Fig. 4D). If the periodicity of this superstructure

matches that of the electrons at the Fermi level, an energy gap will open. These transitions are suppressed in systems where the constituent molecules interact closely between the stacks as well as within the stacks in the crystal lattice. Salts of the electron donor molecule BEDT-TTF (4) with inorganic anions such as I₃⁻, Cu(SCN)₂⁻ and Cu[N(CN)₂]⁻Br⁻, are prime examples: in these CT salts, the metal-insulator instability is eliminated entirely, and superconductivity is observed at temperatures below 8–12 K (ref. 4).

Langmuir-Blodgett films of CT salts

As mentioned previously, crystals of CT salts tend to be brittle and difficult to process, thus limiting their technological potential. Those problems are largely overcome by incorporating the molecules into a Langmuir-Blodgett film structure^{11,12} (Box 1), an approach that appears promising for the molecular engineering of ultra-thin electronic and/or optoelectronic devices. For crystalline materials, the ability to predict the stacking mode of molecules is still very limited, and thus the rational design of conducting crystals of CT salts is essentially restricted to controlling those properties of the component molecules that can be deduced *a priori* (such as planarity, redox potential and extent of π -electron conjugation). But at the intermolecular level, the LB technique has important advantages. Specifically, the individual charge-transfer molecules can adopt only a limited range of packing modes; generally face-to-face and at uniform distances in one-dimensional stacks with a stacking direction that lies parallel to the plane of the substrate. This allows efficient electron delocalization (and conduction) along the stack direction. The central challenge for chemists is therefore the preparation of donor and acceptor molecules that possess excellent film-forming characteristics while satisfying the stringent requirements for high conductivity. The attachment of long chains to render the CT molecules amphiphilic is required for the formation of LB films, but is frequently a difficult task.

The component molecules of conducting LB films^{13,14} are, to date, all derived from molecules that were already known to form conducting CT crystals; notably the donor systems TTF (3) and BEDT-TTF (4), and the acceptor systems TCNQ (1)

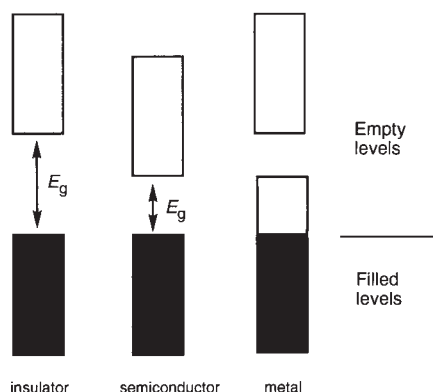


FIG. 1 Band structure of a solid; this determines the electrical properties. E_g is the energy gap between occupied and empty states.

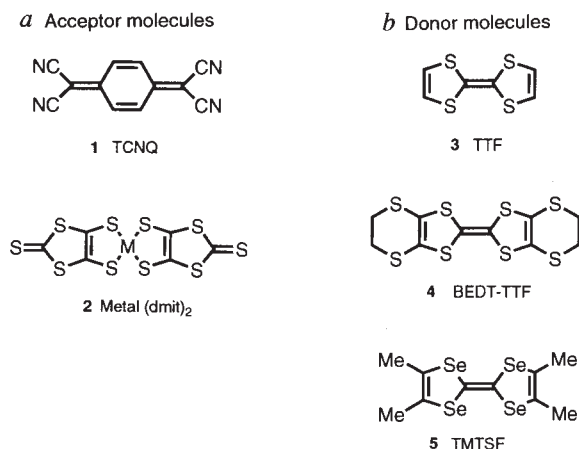


FIG. 2 Formulae of acceptor molecules **1** and **2** (metal bis-(4,5-dimercapto-1,3-dithiole-2-thione), metal(dmit)₂) and donor molecules **3–5** that form crystalline molecular metals. Molecules **2**, **4** and **5** also form superconducting salts. The materials can be single-chain conductors, for example, salts of tetracyano-*p*-quinodimethane (TCNQ, **1**) anion radicals with closed-shell cations for example (pyridinium, sulphonium or alkali metals), or salts of tetrathiafulvalene (TTF, **3**), bis(ethylenedithio)-tetrathiafulvalene (BEDT-TTF, **4**), or tetramethyltetraselenafulvalene (TMTSF, **5**) cation radicals with closed-shell anions (for example, I₃[−], BF₄[−], PF₆[−]), or two-chain conductors, for example, TTF^{•+}–TCNQ^{•−} charge-transfer complexes in which both components are open-shell molecules. The X-ray crystal structures of TTF^{•+}–TCNQ^{•−} and (TMTSF)₂^{•+}PF₆^{•−} are shown in Fig. 3a⁴³ and b⁴⁴, respectively.

and metal(dmit)₂ (**2**). Various strategies have been explored for obtaining electronic conduction in an LB film. The two most widely used methods are as follows. (1) Form the amphiphilic charge-transfer complex in solution, and subsequently assemble an LB film of the complex. If required, the degree of band filling can then be modified by chemical oxidation of the film on the substrate. This usually involves exposure of the film to iodine vapour in a sealed container, to form a tertiary salt of general formula (donor⁺)₁ (acceptor_{*n*})⁻ (I₃)_{1-*n*}. (2) Form an LB film of the neutral donor amphiphile alone, followed by chemical oxidation to the conducting cation radical salt, of general formula (donor⁺)₁ (X⁻)_{*n*}, where *n* < 1 is chosen to give a partially filled band structure. The presence of only one bulky molecular component (the donor) during the film-forming stage results in

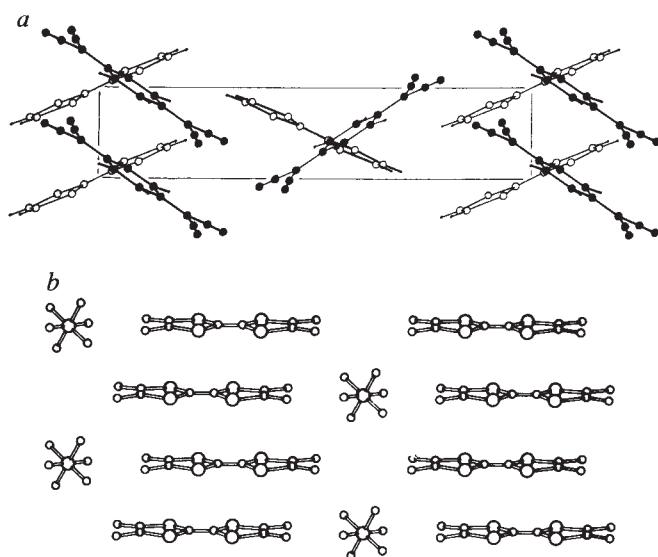


FIG. 3 X-ray crystal structure of charge-transfer salts: a $\text{TTF}^{+}\text{-TCNQ}^{-}$, which is a two-stack conductor (from ref. 43); b, $(\text{TMTSF}_2)^{+}\text{PF}_6^{-}$, which is a single-stack conductor (from ref. 44).

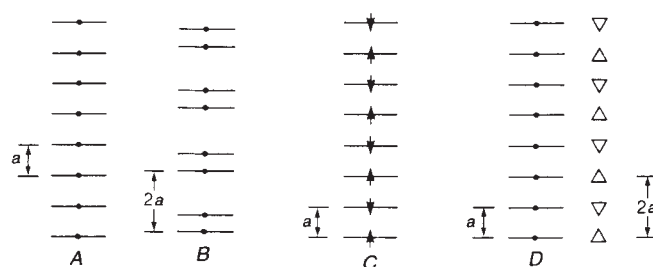


FIG. 4 Several possible configurations of a one-dimensional stack of molecules with one electron in the HOMO: A, metal with uniform lattice constant a ; B, insulator with dimerization resulting from a Peierls transition; C, SDW insulator with spin periodicity caused by Coulombic interactions; D, insulator with periodicity caused by ordering of non-symmetric anions (redrawn from ref. 2).

a more regular film structure than those formed directly from the CT complex. Other methods are also available. For example, the solution of the neutral donor amphiphile can be spread on a subphase containing an oxidant (such as Fe^{3+} or iodine), resulting in the direct formation of a monolayer of the cation radical salt, which may then be deposited onto a substrate as normal. Alternatively, LB films of a neutral donor amphiphile on a conducting support (for example, an electrode coated with gold or indium tin oxide) may be oxidized electrochemically. Electrochemical techniques provide a more controlled means of oxidizing LB films to a conducting mixed-valence state, compared to other chemical methods¹⁵.

To illustrate the progress that has been made to date, we will now review the results from three representative classes of CT complexes that have been shown to form conducting LB films. The molecular formulae of these complexes are given in Fig. 5. The electrical properties of the resulting LB films are summarized in Table I.

TCNQ anion radical salts. Most conducting LB films of CT complexes are based on salts consisting of the TCNQ (**1**) anion radical and a long-chain pyridinium cation. The first films in this class were synthesized from the 1 : 1 (donor:acceptor) charge-transfer complex **6** by Barraud *et al.* (ref. 16). As deposited, films of **6** exhibit low in-plane electrical conductivity, consistent with their having a filled-band structure. The conductivity increases markedly on doping with iodine vapour, but this is not simply a consequence of having changed the occupancy of the electronic bands; the addition of iodine also has a profound effect on the structural organization of the films¹⁷. Before dop-

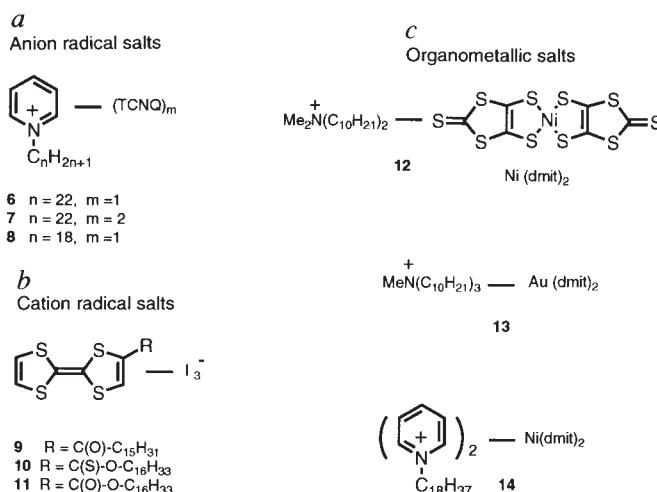


FIG. 5 Structures of amphiphilic CT materials which form conducting LB films.

TABLE 1 Properties of LB films of charge-transfer materials

CT material	Conductivity* as-deposited (S cm ⁻¹)	Conductivity after doping (S cm ⁻¹)	Reference
6	10 ⁻⁵ –10 ⁻⁷	10 ⁻¹ , $E_a \approx 0.1$ eV, (I ₂)	16, 18
7	10 ⁻² , $E_a = 0.3$ eV		19
8	10 ⁻² , $E_a = 0.13$ eV		21
9	10 ⁻⁵	10 ⁻² , $E_a = 0.19$ eV, (I ₂)	24–26
10	4 × 10 ⁻²	1.0, $E_a = 0.09$ eV, (I ₂)	27
11	10 ⁻⁵	10 ⁻² , (I ₂)	42
12		10 ⁻² , (Br ₂)	30
13		25–33 (ClO ₄), 10–15 (Br ₂)	31–33
14	6 × 10 ⁻⁶	0.2–0.8, $E_a = 0.09$ –0.18 (I ₂)	34, 35

The bold numbers in the first column refer to the compounds shown in Fig. 5. The room-temperature values of in-plane conductivity (σ_{in}) and activation energy (E_a) are shown.

* It should be noted that two distinct means of evaluating the lateral conductivity of multilayer films from the measured resistance values exist in the literature. In the first method, used by Barraud and co-workers¹⁶, the calculation neglects the insulating (hydrocarbon) portion of the molecules; in contrast, in the work of Nakamura and co-workers¹⁸ the full length of the substituted charge-transfer system is used. The former technique can produce a conductivity figure up to an order of magnitude than the latter method.

ing, the films consist of sheets of (TCNQ⁻)₂ dimers whose molecular planes are almost parallel to the plane of the substrate. After doping, however, the TCNQ molecules have reoriented so that their molecular planes and long axes are roughly normal to the substrate—as this orientation should improve the overlap between the molecular orbitals, a large enhancement in conductivity is not unexpected. Varying the LB deposition conditions also offers some control over the molecular orientation of this complex¹⁸.

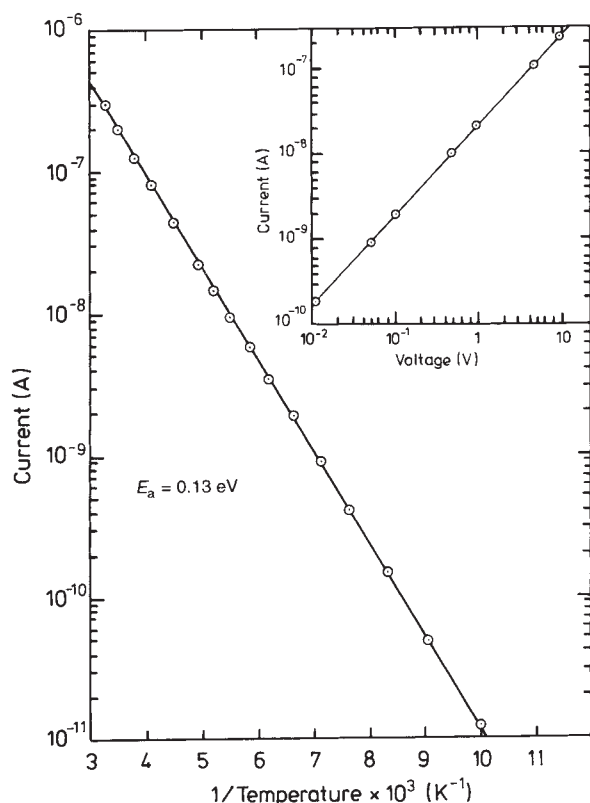


FIG. 6 Semi-logarithmic plot of current versus reciprocal temperature for a nine-layer LB film of TCNQ salt **8** (bias voltage 10 V). Inset, linear current-voltage characteristic obtained for this sample at 20 °C.

Modifying the ratio of donors to acceptors is another route for obtaining a partially filled band, and thus increasing conductivity. This is illustrated by films of the 1 : 2 complex **7**, which exhibit high conductivity without the need for doping or other post-deposition treatment¹⁹. Intriguingly, we have found²⁰ that multilayers of the 1 : 1 complex **8** (which has a filled-band structure) are also conducting, suggesting that they are doped unintentionally during the deposition process (possibly with an anion such as OH⁻).

The conductivity of films based on complex **8** is independent of film thickness and isotropic in the plane of the film. Over the temperature range studied, the conductivity exhibits an exponential dependence on temperature (Fig. 6), characteristic of a semiconductor. The conductivity may therefore be represented by

$$\sigma \propto \exp(-E_a/kT)$$

where E_a is a thermal activation energy. For an intrinsic (undoped) semiconductor, the activation energy is equal to one-half of the energy gap between the valence and conduction bands; for a doped semiconductor, it is related to the energy difference between the dominant dopant level (which usually lies within the band gap) and one of the band edges. For complex **8**, the value of E_a (0.13 eV) is typical of that of a doped inorganic semiconductor, implying a similar conduction process; however, it is still unclear precisely which energy levels control electrical transport in these (and other) films of CT salts.

The structure of films of complex **8**, deduced from ultraviolet-visible dichroism and polarized infrared spectra, are shown in Fig. 7 (ref. 21). In contrast to the orientation of complex **6** (both before and after doping), the TCNQ molecules are inclined at an angle of $\sim 30^\circ$ to the substrate normal.

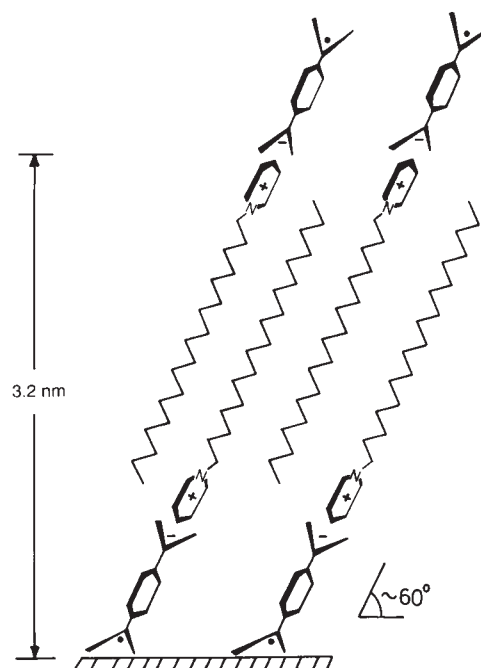
Cation radical salts. LB films based on this class of CT salt have yielded the highest quality films to date, probably as a consequence of the deposition process (the films are first formed from the neutral donor amphiphile, and then chemically oxidized on the substrate to give the CT salt). In most cases^{22,23}, the donor molecules are the amphiphilic derivatives of TTF (**3**), which can be formed by attaching a hydrophobic tail at least 16 atoms in length (required to give a stable monolayer); molecules **9–11** are typical examples.

LB films of compound **9** have been studied extensively^{24–26}. As deposited, the films are insulating, as a charge-transfer complex has yet to be formed. On doping with iodine, the system is first oxidized to a full (1 : 1) charge-transfer state, and is also insulating. But after several hours in air, partial release of iodine leads to a stable, mixed-valence conducting state, and the resulting film has an absorption spectrum characteristic of an organic conductor. Doping is uniform in films consisting of at least 40 layers, and does not significantly impair the integrity of the structure. The molecules appear to adopt a near-vertical orientation with respect to the substrate, and there is some evidence for interdigitation after doping²⁵.

Similar structural ordering and doping characteristics are observed for LB films of compound **10** (ref. 27), but with the important difference that the in-plane room-temperature conductivity is two orders of magnitude higher after doping. This is the highest conductivity yet achieved for a TTF-based LB film, but is still well below that of single-crystal TTF-halide salts (for example²⁸, TTF-Br_{0.77} has a room-temperature conductivity of ~ 800 S cm⁻¹). The effects of iodine doping can be monitored by infrared spectroscopy^{27,29} (Fig. 8). It seems clear that the C=S group is important for the high conductivity of compound **10** after doping, but its precise role is not yet understood.

Organometallic CT salts. Crystalline CT salts based on metal(dmit)₂ anions **2** have yielded both metallic and superconducting materials^{6,7}. These anions can be used in conjunction with both radical (open-shell) cations such as TTF (**3**) and closed-shell cations such as tetraalkylammonium and sodium⁷. Furthermore, the peripheral sulphur atoms of the dmit system can play an important role in both intra- and inter-stack inter-

FIG. 7 Schematic diagram showing the proposed average orientation of salt **8** in LB film form. Ultraviolet-visible dichroism and polarized infrared spectra indicate that in films of salt **8**, the TCNQ molecules and the alkyl chains have their long axes inclined at an angle of $\sim 30^\circ$ to the substrate normal. From low-angle X-ray diffraction data the Bragg d -spacing was found to be 3.2 nm; taken together these data are consistent with the interdigitated multilayer structure.



actions, which helps to suppress transitions to an insulating state at low temperatures. Metal(dmit)₂ anions have been successfully incorporated into LB films by complexing them with long-chain ammonium or pyridinium cations, **12–14**.

The first conducting films to be formed by this route were based on bromine-doped LB films of complex **12** (ref. 30). Although this gives reasonable levels of conductivity, the film-forming characteristics of the complex are poor. LB films based on complex **13** (combined with a fatty acid) appear more promising^{31–33}. After chemical or electrochemical oxidation with perchlorate ions, the films exhibit high room-temperature conductivity; moreover, metallic (rather than semiconducting) behaviour is observed down to ~ 200 K. The conductivity data for these films are consistent with a two-dimensional variable-range hopping model, in which polycrystalline regions of high conductivity are separated by material of low conductivity^{32,33}; improving the long-range order in these films is therefore expected to lead to further increases in conductivity.

Complexes **12** and **13** both have 1:1 stoichiometries, and therefore need to be doped to achieve high conductivities. The 2:1 charge-transfer complex **14** can also form LB films, with

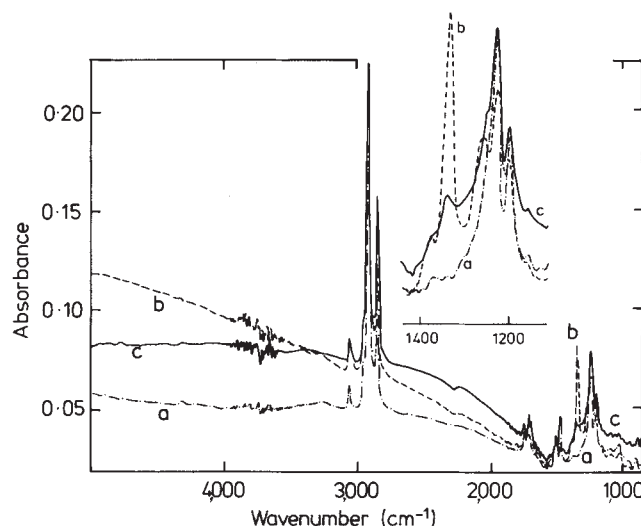
the important advantage that addition of a fatty acid is not required to form a stable monolayer at the air–water interface³⁴. As deposited, the room-temperature conductivity of this complex is low, but increases dramatically after exposure of the films to iodine vapour^{34–36} (the complexes appear to be oxidized in the vicinity of the nickel centre). True metallic behaviour has yet to be observed in films of this complex, and it seems clear that the processes responsible for its electrical properties are distinct from those in the TCNQ- and TTF-based films.

The long-term stability of the electrical properties is an important consideration for any device application, but is often overlooked. It is therefore worth noting that multilayer structures of complex **14** show some of the best stability characteristics to date; we have found that samples stored in air (but in the absence of light) show no significant change in conductivity over at least two months³⁵.

Potential device applications

There have been several attempts to incorporate LB layers into electronic device structures¹¹. Most of the early work concerned the deposition of fatty-acid layers onto a semiconductor, to pas-

FIG. 8 Transmission infrared spectra of LB films of compound **10** (96 layers) including an expansion of the 1,400–1,100 cm⁻¹ region: trace a, before doping; trace b, 5 min after doping; trace c, 2 h after doping. Before doping, the C=S band is present at 1250 cm⁻¹ (trace a); this band shifts immediately upon doping, consistent with oxidation having occurred, i.e. neutral **10** → radical cation **10**. The new sharp band which appears at 1350 cm⁻¹ is a vibronic coupling band, and a broad CT band is now seen at $\sim 5,000$ cm⁻¹ (trace b). Two hours after doping, when the conductivity value has reached its maximum, the CT band has shifted to lower wavenumber and the vibronic coupling band is much weaker (trace c). The role of the C=S group of compound **10** in improving the conduction properties of the LB films (compared to analogues **9** and **11**) is not yet understood.



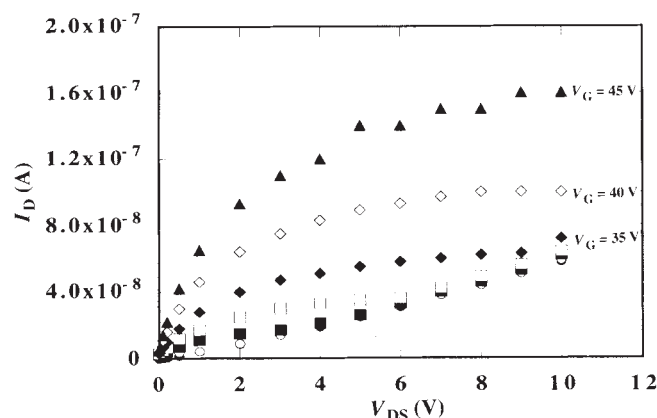


FIG. 9 Drain current I_D versus drain-source voltage V_{DS} at various gate voltages V_G for a thin-film FET incorporating an LB film of compound **14**.

sivate the surface or to modify the electronic band structure at the semiconductor surface. The role of the LB layer was essentially a passive one, providing an insulating layer across which a voltage could be sustained, thereby 'decoupling' the Fermi level of the semiconductor from that of the metal top electrode. In this way, accumulation, depletion and inversion of charge carriers in the semiconductor can be achieved. Field-effect transistors (FETs) have now been demonstrated with a few LB-film/semiconductor combinations¹¹.

The introduction of the new semiconducting or conducting multilayers of charge-transfer complexes opens up a wide range of other possibilities for LB-film electronic devices³⁷. One of the most important developments in this area has been the incorporation of an LB film as the active semiconductor layer in an FET. Such structures have been prepared using the soluble conjugated polymer poly(3-hexylthiophene) (ref. 38). We have fabricated FETs using the charge-transfer complex **14**. Figure 9 shows a typical set of current versus voltage characteristics³⁶ for such a device, which are consistent with standard transistor theory. The carrier mobilities for both organometallic and polymer FETs

are, however, quite small. The mobility for the organometallic device can be increased by several orders of magnitude, to $\sim 0.2 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ by doping the LB layers with iodine. These mobilities are comparable to those found for the early thin-film transistors based on amorphous silicon. Of course, devices based on LB films will never compete with the current generation of inorganic devices based on single crystals or quantum wells; but if they can be made sufficiently reliable, they may find a niche in areas such as the addressing of large-area displays, or as interfaces to other organic or biological films in highly specific chemical sensors. Perhaps the most exciting prospects for these films lie in the ability to tailor their electrical properties at both the synthesis and deposition stage. By mixing different molecules within each LB layer, or by building up multilayer structures of different layers, new materials or devices may be engineered at the molecular level.

Much more work is needed before it will be clear whether LB films of CT salts can find their way into the commercial world. LB films are not single crystals, and a deeper understanding of the nature of their charge-transport processes is required. The high in-plane conductivities achieved to date are promising from a practical point of view, but are still likely to be limited by poor electrical transport between polycrystalline domains within the films; methods for improving long-range structural order are needed. Long-term stability is also a problem, one solution of which may involve incorporation of the charge-transfer complexes into more robust polymeric LB films.

The amphiphilic nature of the LB films also presents a problem. The aliphatic chains, essential for providing insoluble and ordered monolayers, also act as rather good electrical insulators; this can be a serious impediment in applications where good conductivity perpendicular to the plane of the film is required. A number of groups, including our own, are exploring the possibility of incorporating conducting 'molecular wires' into multilayer structures, to enhance the electrical connectivity between the conducting planes. Polyene molecules based on amphiphilic carotenoids appear particularly promising for this purpose^{39, 41}.

We have considered just a few of the many possibilities for using the LB technique to engineer novel organic multilayer structures based on charge-transfer materials. Ultimately, the success of this field—scientifically or technologically—will depend on the ability to produce well defined and well characterized ultra-thin films of pure and stable materials. □

- Bryce, M. R. Murphy, L. C. *Nature* **309**, 119–126 (1984).
- Greene, R. L. & Street, G. B. *Science* **226**, 651–656 (1984).
- Cowan, D. O., Fortkrot, J. A. & Metzger, R. M. in *Lower-Dimensional Systems and Molecular Electronics* (ed. Metzger, R. M.) 1–22 (Plenum, New York, 1991).
- Williams, J. M. et al. *Science* **252**, 1509–1514 (1991).
- Jérome, D. *Science* **252**, 1509–1514 (1991).
- Ishiguro, T. & Yamaji, K. *Organic Superconductors* (Springer, Berlin, 1990).
- Williams, J. M. et al. *Organic Superconductors (Including Fullerenes)* (Prentice-Hall, New Jersey, 1992).
- Proc. ICSM 92 in *Synth. Metals* **55–57** (1993).
- Fröhlich, H. *Proc. R. Soc. A* **223**, 296 (1954).
- Peierls, R. e. *Quantum Theory of Solids* (Oxford Univ. Press, London, 1955).
- Roberts, G. G. (ed.) *Langmuir–Blodgett Films* (Plenum, New York, 1990).
- Kenn, R. M. et al. *J. phys. Chem.* **95**, 2092–2097 (1991).
- Nakamura, T. & Kawabata, Y. *Techno Japan* **22**, 8–17 (1989).
- Tieke, B. *Adv. Mater.* **2**, 222–231 (1991).
- Goldenberg, L. M. et al. *Thin Solid Films* **238**, 280–284 (1994).
- Barraud, A., Ruauadel-Teixier, A., Vandevyver, M. & Lesieur, P. *Nouv. J. Chim.* **9**, 365–367 (1985).
- Richard, J. et al. *J. chem. Phys.* **86**, 2428–2438 (1987).
- Nakamura, T., Tanaka, M., Sekiguchi, T. & Kawabata, Y. *J. Am. chem. Soc.* **108**, 1302–1303 (1986).
- Matsumoto, M. et al. *Synth. Metals* **19**, 675–680 (1987).
- Ward, R. J., Dhindsa, A. S., Bryce, M. R., Petty, M. C. & Monro, H. S. *Thin Solid Films* **198**, 363–367 (1991).
- Dhindsa, A. S. et al. *J. molec. Electron.* **5**, 135–142 (1989).
- Bryce, M. R. et al. *J. chem. Soc., chem. Commun.* 816–818 (1990).
- Batsanov, A. S. et al. *Chem. Mater.* **6**, 1419–1425 (1994).
- Dhindsa, A. S., Bryce, M. R., Lloyd, A. S. & Petty, M. C. *Thin Solid Films* **165**, L97–L100 (1988).
- Dhindsa, A. S. et al. *Synth. Metals* **35**, 307–318 (1990).
- Dhindsa, A. S., Bryce, M. R., Ancelin, H., Petty, M. C. & Yarwood, J. *Langmuir* **6**, 1680–1682 (1990).
- Dhindsa, A. S. et al. *Chem. Mater.* **4**, 724–728 (1992).

- Scott, B. A., Placa, S. J., Torrance, J. B., Silverman, B. D. & Welber, B. J. *Am. chem. Soc.* **99**, 6631–6639 (1977).
- Song, Y. P., Dhindsa, Bryce, M. R., Petty, M. C. & Yarwood, J. *Thin Solid Films* **210/211**, 589–591 (1992).
- Nakamura, T., et al. *Chem. Lett.* 1667–1670 (1988).
- Nakamura, T. et al. *Chem. Lett.* 367–370 (1989).
- Nakamura, T., Miura, M., Matsumoto, M., Tachibana, H., Tanaka, M. & Kawabata, Y. in *The Physics and Chemistry of Organic Superconductors* (eds. Saito, G. & Kagoshima, S.) 424–427 (Springer Proc. Phys. Vol. 51, Springer, Berlin, 1990).
- Miura, Y. F. et al. *Jap. J. appl. Phys.* **30**, 3503–3509 (1991).
- Dhindsa, A. S., Badyal, J. P., Pearson, C., Bryce, M. R. & Petty, M. C. *J. chem. Soc., chem. Commun.* 322–323 (1991).
- Pearson, C., Dhindsa, A. S., Petty, M. C. & Bryce, M. R. *Thin Solid Films* **210/211**, 257–260 (1992).
- Pearson, C., Moore, A. J., Gibson, J. E., Petty, M. C. & Bryce, M. R. *Thin Solid Films* **244**, 932–935 (1994).
- Petty, M. C. *Thin Solid Films* **210/211**, 417–426 (1992).
- Paloheimo, J., Stubbs, H., Yli-Lahti, P., Dyreklev, P. & Ingañäs, O. *Thin Solid Films* **210/211**, 283–286 (1992).
- Ohnishi, T., Hatakeyama, M., Yamamoto, N. & Tsubomura, H. *Bull. chem. Soc. Japan* **51**, 1714–1716 (1978).
- Wegmann, A., Tieke, B., Pfeiffer, J. & Hilti, B. *J. chem. Soc. chem. Commun.* 586–588 (1989).
- Williams, G., Bryce, M. R. & Petty, M. C. *Molec. Cryst. Liq. Cryst.* **229**, 83–90 (1993).
- Cooke, G. et al. *Synth. Metals* **55–57**, 3871–3878 (1993).
- Kistenmacher, T. J., Phillips, T. E. & Cowan, D. O. *Acta crystallogr. B* **30**, 763–768 (1974).
- Thorup, N., Rindorf, G., Soling, H. & Bechgaard, K. *Acta crystallogr. B* **37**, 1236–1240 (1981).

ACKNOWLEDGEMENTS. The work at Durham University on conducting LB films has been supported by the EPSRC. M. R. Bryce is at the Department of Chemistry, and M. C. Petty is at the School of Engineering, University of Durham, Durham, DH1 3LE, UK. The authors are co-directors of the University of Durham Centre for Molecular Electronics.

In vitro evolution of a self-alkylating ribozyme

Charles Wilson* & Jack W. Szostak

Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA

RNA enzymes are postulated to have catalysed all chemical reactions in the earliest living cells. This idea is now investigated in a search for alkyl transferases from a pool of random sequence RNAs. Selection for self-biotinylation yields a transfer RNA-like ribozyme that efficiently catalyses carbon–nitrogen bond formation. Ribozymes can thus promote reactions other than those involving the RNA sugar–phosphate backbone, suggesting that RNA may be capable of a broad range of catalytic activities.

BOTH the genetic and enzymatic components of the earliest cells are thought to have been RNA molecules, because RNA is the only known macromolecule that can both encode information in a heritable form and act as a biocatalyst¹. This basic 'RNA world' model has been extended with the proposal that much of modern metabolism evolved before the evolution of encoded protein synthesis, and that early ribozyme-catalysed metabolic transformations form the basis of our present protein-catalysed metabolism². This proposal requires that ribozymes should be able to catalyse a broad range of chemical transformations. However, all known natural ribozymes, including the group I and group II introns, RNase P, and the hammerhead and hairpin RNAs, catalyse a restricted range of reactions involving the RNA sugar–phosphate backbone³. It is not known whether this restriction to a set of hydrolysis and transesterification reactions is an inherent limitation of the catalytic ability of RNA, or whether protein enzymes have simply outperformed their RNA ancestors and thus displaced them in the course of evolution. To address this question, we have used *in vitro* selection to search through pools of random sequence RNAs for enzyme-like molecules, attempting to recover new types of ribozymes similar to those that may have functioned in the earliest living cells.

Relatively short RNAs (aptamers) capable of binding with high affinity and specificity to a variety of small ligands have been isolated from pools of 10^{13} to 10^{14} different RNA sequences^{4–10}. In addition, ribozymes that act as oligonucleotide ligases, polynucleotide kinases, and isomerases have been obtained from similar pools^{11–13}. Here we describe the isolation, from pools of random sequence molecules, of RNAs capable of binding to non-nucleotide substrates and rapidly performing an alkylation reaction. The successful isolation of self-alkylating ribozymes is of general interest for two main reasons. First, transfer RNAs, ribosomal RNAs and messenger RNAs are generally modified post-transcriptionally by well-conserved alkylations, many of which are required for biological function. Our current findings raise the possibility that these modifications may have initially been introduced as RNA-catalysed events. Second, alkylation typically occurs by an S_N2 mechanism and, as such, the mechanisms used to promote alkylations are the same as those used for many other biologically important transformations, including hydrolysis and phosphoryl and acyl transfers. The isolation of a highly active self-alkylating RNA suggests that ribozymes may indeed be able to catalyse a wide range of reactions, and may have been important in the early evolution of metabolism.

Strategy for evolving new ribozymes

Our strategy for isolating new ribozymes involved a series of sequential *in vitro* selections (Fig. 1). We reasoned that the num-

ber of functional ribozymes in a pool of completely random RNAs might be vanishingly small, as catalysis of a complex reaction would demand not only the ability to bind a non-RNA substrate, but also the preferential stabilization of the transition-state configuration of reactants. We therefore designed our experiment as a three-stage selection, searching initially for RNAs capable of binding the desired substrate, but not requiring that these RNAs also possess catalytic activity. After isolating such a substrate-binding RNA by repeated rounds of affinity chromatography and amplification, we generated a second pool of RNAs which contained the selected binding domain sequence extensively modified by randomly distributed base substitutions. We anticipated that some small fraction of these new RNAs might contain the appropriate set of mutations that would introduce enzyme-like activity without disrupting substrate binding. Relatively inefficient ribozymes were isolated at this stage, and were subsequently optimized by a third additional step of *in vitro* evolution.

In the first application of this strategy, we attempted to isolate a ribozyme that performs a covalent, self-labelling reaction with a biotin derivative. Ribozymes with this activity can be readily detected and enriched *in vitro* because biotinylated RNAs can be efficiently trapped by agarose-immobilized streptavidin. The extremely high affinity of streptavidin for biotin ($K_A \approx 10^{15} \text{ M}^{-1}$) allows extensive washing under RNA-denaturing conditions to remove virtually all non-biotinylated RNAs, leaving only labelled molecules for subsequent amplification¹⁴. Here we describe an application of this procedure to the isolation of RNAs that carry out a specific alkylation reaction. In combination with the appropriate biotin derivative and end-modified RNA pool, the same process could be used to select for ribozymes that perform any type of chemical condensation reaction (for example, peptide bond formation, glycosylation, and so on).

Isolation of biotin-binding RNAs

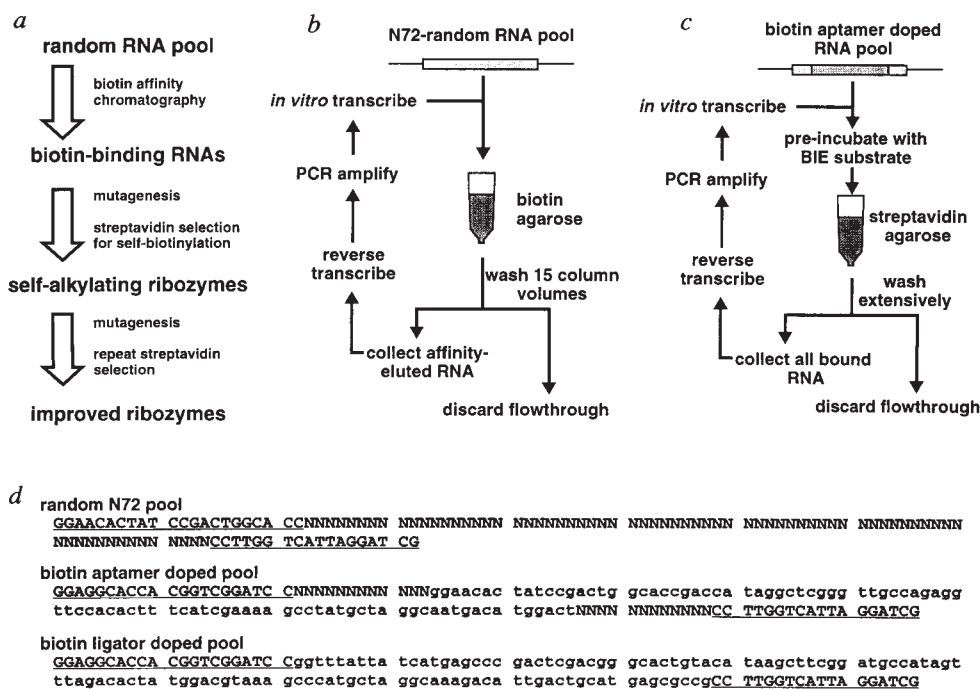
A pool of approximately 5×10^{14} different random sequence RNAs was generated by *in vitro* transcription (Fig. 1*d*). On average, any given 28-nucleotide sequence has a 50% probability of being represented in a pool of this complexity. Biotin-binding RNAs were isolated by affinity chromatography with a biotin agarose matrix, carried out as described previously and in the legend for Fig. 1. Reverse transcription of the specifically eluted RNA, followed by PCR amplification and *in vitro* transcription, yielded an enriched pool of RNA for subsequent reselection.

After seven rounds of repeated enrichment, more than half of the RNA pool specifically binds to biotin (Fig. 2*a*). The pool of DNA templates corresponding to molecules isolated after seven rounds of selection was cloned and individual aptamers were sequenced. Remarkably, a single sequence (represented by clone BB8-5) accounted for more than 90% of the selected pool (two minor clones account for most of the remaining RNAs). This result indicates that there are relatively few solutions to the prob-

* Present address: Department of Biology and Center for the Molecular Biology of RNA, University of California, Santa Cruz, California 95064, USA.

FIG. 1 *a*, Strategy for *in vitro* evolution of self-alkylating ribozymes. *b*, Scheme for isolating biotin-binding RNAs by affinity chromatography. *c*, Scheme for isolating self-biotinylating RNA enzymes. *d*, Coding sequences for RNA pools used for *in vitro* selection experiments. Upper case A, C, G and T represent pure nucleotide; N represents an equimolar mix of A, C, G and T; lower case a, c, g and t represent 70% of one nucleotide with 10% of the other three nucleotides; underlined regions, constant primer sequences used for amplification.

METHODS. Random sequence RNA (approximately 80 μ g in the first round, corresponding to 2–3 copies of each sequence; approximately 10 μ g thereafter) was resuspended in wash buffer (100 mM LCI, 10 mM Na-HEPES, pH 7.4, 5 mM MgCl₂), applied to a 0.5-ml biotin agarose, and washed with 15 column volumes of wash buffer (1 column volume, 500 μ l). After affinity elution of the biotin-binding RNA with 5 column volumes of buffer containing 5 mM biotin, 10 μ g glycogen and 0.3 M NaCl were added and the mixture was precipitated with 2.5 volumes of ethanol at -78°C . After resuspending the selected RNA, the reverse transcriptase primer (2.5 μ M) was annealed at 65°C for 3 min, and reverse transcription was carried out at 42°C for 45 min (using Superscript RT enzyme; Life Technologies). PCR amplification was performed by diluting one-fifth of the RT reaction with the appropriate dNTPs, PCR buffer, USB Taq Polymerase, and 0.5 μ M(+) primer containing the T7 RNA polymerase promoter (in the first round, when the number of copies of each RNA sequence was small, all of the reverse transcription reaction was amplified). A strong band of the correct size was typically observed after 8–15 cycles amplification (94°C , 1 min; 55°C , 45 s; 72°C , 1 min). Half of the PCR reaction was used for *in vitro* transcription with T7 RNA polymerase (37°C , overnight). The resulting RNA was purified by electrophoresis on an 8% polyacrylamide gel and used for the subsequent round of selection. The RNA pool derived from the BB8-5 sequence was co-incubated for a defined time with 200 μ M BIE (Molecular Probes) under folding conditions (100 mM LCI, 10 mM Na-HEPES, pH 7.4, 5 mM MgCl₂). Reaction with BIE was terminated by the addition of 100 mM β -mercaptoethanol, 5 mM EDTA, 0.3 M NaCl, 50 μ g tRNA (*Escherichia coli*, RNase-free, Boehringer). After 5 min the mixture was precipitated with 2.5 volumes ethanol. After washing and resuspension, the RNA was applied to 0.5 ml of a 50% slurry of streptavidin agarose in wash buffer (1 M NaCl, 10 mM Na-HEPES, pH 7.4, 5 mM EDTA) that had been prewashed with 50 μ g tRNA. After rocking for 30 min to allow



streptavidin-biotin binding, the mixture was transferred to a 10-ml column and washed with 4×10 ml wash buffer and 2×10 ml distilled water. Specifically bound RNAs were eluted by heating in the presence of 10 mM biotin at 95°C for 8 min. Amplification of the selected RNAs was achieved as described above for the biotin binder selection. The random N72 pool was designed and synthesized by D. P. Bartel as described previously (ref. 11). For the biotin aptamer doped pool, deletion analysis indicated that the 3' primer was not required for binding and the same sequence was therefore used for the 3' primer of the partially randomized pool. To allow for the possibility that the 5' primer formed part of the aptamer core, the original 5' primer sequence was included in the partially randomized region of the new pool and a different 5' primer was appended for amplification. Because of differences in the relative rates of phosphoramidite incorporation during DNA synthesis, a biased mix of all four nucleotides was prepared with molar ratios of 3:3:2:2 (A:C:G:T). This mix was added to pure phosphoramidite stocks (a and c, 64% pure stock, 36% random mix; G and T, 55% pure stock, 45% random mix) to yield mixed stocks for pool synthesis. Synthesis yielded 77 μ g DNA (1.52 nmol). Primer extension indicated that only 8.7% of the DNA molecules could serve as full-length templates for Taq polymerase. The pool thus contains $1.52 \times 10^{-9} \times 6.02 \times 10^{23} \times 0.087 = 8 \times 10^{13}$ distinct sequences. The biotin ligator doped pool was prepared exactly as described above for the biotin aptamer doped pool, but using the sequence of the BL8-6 clone.

lem of binding biotin with the required affinity in the sequence space spanned by the RNA random pool. Previous RNA selection for binding to small ligands, including various dyes, amino acids, cofactors and nucleotides, have suggested that aptamers exist at a frequency of 10^{-10} to 10^{-11} in random sequence pools^{3,4}. All of these ligands, however, have contained aromatic rings which could intercalate between RNA bases, and/or charged groups which might interact electrostatically with the RNA backbone. The significantly lower frequency of biotin binders ($\sim 10^{-14}$) shows that ligands lacking such groups may pose more difficult recognition problems for RNA, leading to a requirement for a more complex binding site.

Selecting for biotin-utilizing ribozymes

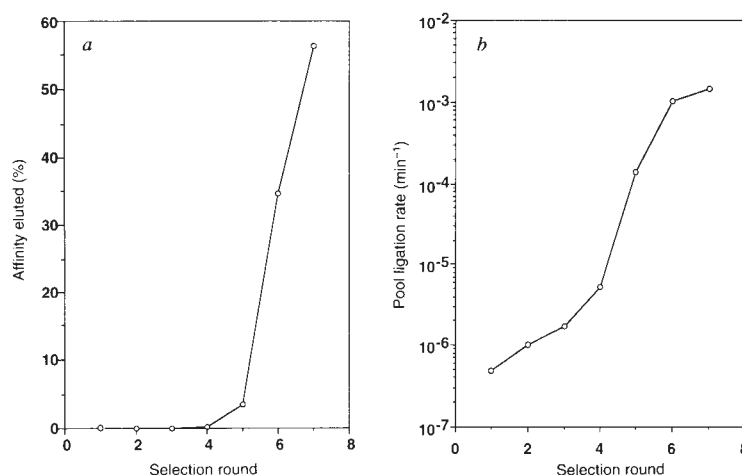
The sequence of the BB8-5 biotin aptamer was used to direct the synthesis of a second pool of RNAs which was screened for the presence of biotin-utilizing ribozymes. This pool contained a core of 93 nucleotides specified by the aptamer sequence with the wild-type nucleotide incorporated at each position in the template with 70% probability (the three non-native nucleotides

each occurring with 10% probability) (Fig. 1d). Transcription yielded a pool of 8×10^{13} unique RNA sequences clustered in sequence space around the original biotin aptamer sequence.

Using this second RNA pool, we searched for ribozymes able to enhance the rate of self-alkylation with the haloacetyl derivative, *N*-biotinoyl-*N*-iodoacetyl-ethylenediamine (BIE). BIE can be used to biotinylate proteins by reaction with free cysteine sulphhydryls¹⁵. To provide one potential internal substrate for the alkylation reaction, the doped pool was transcribed in the presence of excess 8-mercaptoguanosine, thus yielding RNAs containing a single free thiol in the 5'-terminal nucleotide. Active ribozymes were isolated by co-incubating the RNA pool with low concentrations of the BIE substrate for an extended time to allow self-modification to occur. RNAs that had become biotinylated were isolated by streptavidin agarose chromatography (Fig. 1c). Covalent attachment of the BIE substrate (as opposed to tight non-covalent binding to either the non-biotin half of BIE or to streptavidin) was ensured by washing the streptavidin column under RNA-denaturing conditions. Specifically bound RNAs could be eluted by extensive heating; amplification of the

FIG. 2 a, Progress of the biotin aptamer selection. Biotin-eluted RNA expressed as a percentage of total RNA applied to the biotin-agarose column is plotted as a function of selection cycle. b, Progress of the self-biotinylation selection.

METHODS. Individual RNAs eluted from the seventh round were subcloned and sequenced; more than 90% of the clones correspond to the sequence shown in Fig. 5. The rate of self-biotinylation was determined by a timecourse experiment. ^{32}P -labelled RNA was first resuspended in incubation buffer (100 mM KCl, 10 mM Na-HEPES, pH 7.4, 5 mM MgCl_2) and allowed to equilibrate for 10 min at room temperature. BIE (200 μM) was added to the mixture and portions were subsequently removed after 0 to 120 min of incubation. Samples were quenched and affinity-purified as described for Fig. 1. Portions were counted in a scintillation counter following ethanol precipitation (total RNA count) and following binding to streptavidin agarose (product RNA count); the ratio of these two counts is the fraction reacted. Values are corrected for 0.02% nonspecific RNA binding.



resultant molecules (by reverse transcription, PCR and transcription) yielded a pool enriched for ribozymes.

After four rounds of selection, an increase in the proportion of RNAs binding to the streptavidin was observed (Fig. 2b). By the fifth round, 10% of the RNA ligated the biotin substrate. To select for the most active ribozymes, the incubation time was progressively shortened from 15 h to 30 min to 1 min. After seven rounds of selection, no further increase in activity was observed, suggesting that the complexity of the starting pool had been exhausted. Sequencing individual clones from the selected pool showed that approximately 50% of the ribozymes were very closely related and were derived from a single progenitor. One of these clones, BL8-6, performs the self-biotinylation reaction at a rate of 0.001 min⁻¹ in the presence of 200 μM BIE.

Optimizing activity

It is likely that the original RNA pool from which the BL8-6 ribozyme derives does not saturate the space of biotin-ligating ribozymes. To test the possibility that appropriate additional mutations to the BL8-6 sequence might increase its activity, a third RNA pool was generated based on its sequence but with non-wild-type nucleotides substituted at each position with 30% probability (Fig. 1). The selection for alkylation activity was repeated as described above, but with both the reaction incubation time and the BIE concentration progressively lowered to select for the most active enzymes. After eight rounds of selection (ending with an incubation period of 1 min at 10 μM BIE), active clones from the pool were sequenced and assayed for activity. Ribozymes in this pool were uniformly more active than their BL8-6 progenitor, with one clone (BL2.8-7) self-biotinylation at a rate of 0.05 min⁻¹ in the presence of 100 μM BIE (100-fold more active than BL8-6).

Nature of the reaction product

We had originally anticipated that the free thiol in the 8-mercaptoguanosine base at the 5'-end of the RNA might serve as a nucleophile for the alkylation reaction. However, the observation that BL8-6 ribozyme transcribed without 8-mercaptoguanosine performed the self-biotinylation reaction as efficiently as the thiol-containing RNA indicated that some other site was being alkylated. To identify the reactive site, 5'-end labelled BL8-6 ribozyme that had reacted with BIE was subjected to alkaline hydrolysis, and the resultant ladder of molecules was affinity purified on streptavidin agarose. Full-length RNAs and those with up to approximately 60 3'-terminal nucleotides deleted were retained by the streptavidin, whereas shorter molecules were not bound. This result maps the biotin attachment site to the region ... 5'-⁹²GGACGUAAA¹⁰⁰-3' ... Alkylation at the N7 position of purines leads to RNA strand scission after treatment with sodium borohydride followed by aniline acetate (this reaction serves as the basis for RNA chemical sequencing)¹⁶. RNA incu-

bated with BIE, purified on streptavidin-agarose and treated in this manner was cleaved at G⁹⁶ (... GGACGUAA ...) (Fig. 3a). No G⁹⁶-specific cleavage was observed for RNA that had been exposed to BIE but not biotinylated (that is, the streptavidin flowthrough fraction). G⁹⁶ is therefore the alkylation site for the ribozyme.

To characterize further the alkylation product, body-labelled biotinylated BL8-6 ribozyme was digested to 5'-monophosphate nucleotides using phosphodiesterase (Fig. 3b). Thin-layer ion-exchange chromatography (TLC) indicated the presence of a radioactive species in the streptavidin-purified RNA that was absent from the streptavidin-flowthrough RNA. This adduct migrated more rapidly than 5'-GMP, with a mobility similar to that for 7-methyl GMP. These results suggest that the adduct carries a positive charge, consistent with alkylation at N7. Although we cannot rule out the possibility of alkylation at N1 or N3, alkylation at either of these sites would not be expected to lead to strand cleavage following aniline treatment, but would be expected to disrupt reverse transcription, thus preventing ribozymes using these nucleophiles from being enriched during the *in vitro* selection procedure. Taken together, these results strongly suggest that N7 of G⁹⁶ is the alkylation site.

The reaction rate enhancement

How good are the selected ribozymes in comparison with typical protein enzymes? To answer this question, the background rate of guanosine alkylation by BIE was determined by two independent methods. First, radiolabelled random sequence RNA (from the pool used to isolate the original biotin binder) was incubated for 24 h with or without 200 μM BIE. The specific increase in the fraction bound by streptavidin agarose (0.15%) after extensive washing was taken as a measure of the background reaction. Assuming an average of 28 guanines per RNA sequence, this fraction corresponds to a non-enhanced alkylation rate of $2.3 \times 10^{-6} \text{ s}^{-1} \text{ M}^{-1}$. In a similar approach, low concentrations of [α - ^{32}P]GTP were incubated overnight in the presence or absence of 200 μM BIE and, after 12 h, affinity purified by streptavidin agarose. The fraction specifically bound (3.4×10^{-5}) indicates a background rate of $2.3 \times 10^{-6} \text{ s}^{-1} \text{ M}^{-1}$, in close agreement with that obtained from the RNA labelling experiment. A timecourse experiment with BL2.8-7 RNA yields a self-biotinylation rate of approximately $8 \text{ s}^{-1} \text{ M}^{-1}$. The ribozyme rate enhancement is thus approximately 3×10^6 , comparable to that of highly active catalytic antibodies (assayed under multiple turnover conditions), although substantially less than that of many natural protein enzymes^{17,18}. It should be noted, however, that the experimental estimate for the background reaction rate includes the sum of all possible alkylation reactions and is not limited to attack by guanosine N7 nucleophiles. As such, it would tend to underestimate the true rate enhancement for alkylation specifically at the guanosine N7 position.

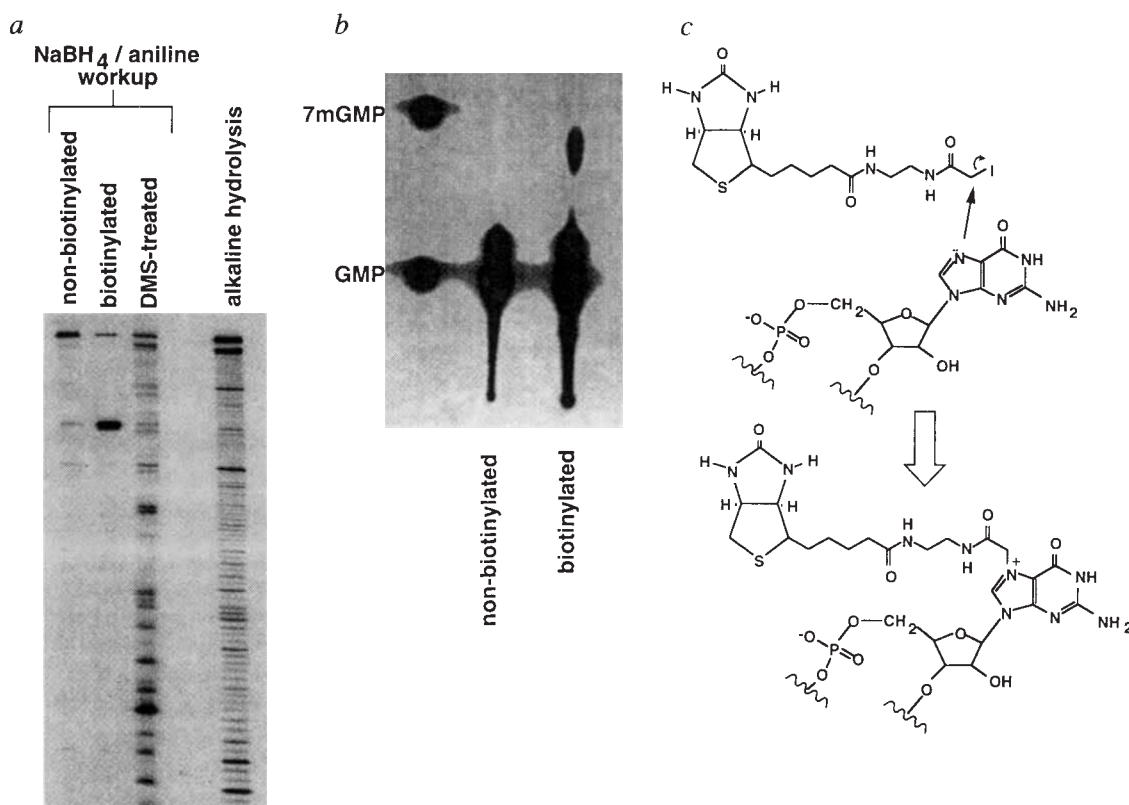


FIG. 3 Site-specific alkylation reaction performed by the BL8-6 ribozyme. *a*, Comparison of biotinylated, non-biotinylated and DMS-treated (control) RNA. The major cleavage site in biotinylated RNA (G⁹⁶) is absent from the non-biotinylated RNA. The control lacks a major band, indicating that the active site nucleotide is not generally activated for reaction with alkylating agents. By comparison to the partial alkaline hydrolysis ladder, the minor bands in the non-biotinylated fraction appear to result from nonspecific RNA cleavage. *b*, TLC characterization of the alkylation product. *c*, Schematic representation of the N-alkylation reaction at the N7 position of G-96.

METHODS. BL8-6 RNA (5' end-labelled) was allowed to react overnight with 200 μ M BIE and then separated by streptavidin affinity chromatography into biotinylated and non-biotinylated fractions. RNA was then treated with sodium borohydride and aniline acetate to cleave

specifically at N7 alkylation sites (RNA was dissolved in 1.0 M Tris-HCl, pH 8.2, and 0.2 M NaBH₄; following a 30-min incubation, the reaction was quenched with 0.6 M sodium acetate/0.6 M acetic acid, pH 4.5, containing carrier tRNA, precipitated and rinsed, and then treated with 1.0 M aniline/acetate, pH 4.5 at 60 °C for 20 min). For comparison, DMS-treated RNA was treated in parallel (RNA was partially hydrolysed by heating to 90 °C for 7 min in the presence of 100 mM NaHCO₃, pH 9.0 and subsequently ethanol precipitated). Biotinylated and non-biotinylated BL8-6 RNA body-labelled with α -³²P-GTP was digested to completion with snake venom phosphodiesterase (RNA was diluted with 25 μ l 10 mM NaCl, 10 mM MgCl₂, 10 mM Tris-Cl, pH 7.4, and 5 μ l phosphodiesterase I (Boehringer; 3 U per ml) and incubated for 20 h at 37 °C. PEI cellulose TLC plates from J. T. Baker Co. were spotted with the digestion reactions, prerun with H₂O to remove excess salts, and then developed with 6 M formic acid. Non-labelled 7-methyl-5'-GMP (7mGMP) and 5'-GMP were run in parallel as standards. After developing, the positions of 7mGMP and GMP were determined by UV shadowing and marked with radioactive tracer.

Three simplistic mechanisms can be proposed to account for the alkylation activity of the selected ribozymes. The reaction might be driven by very tight binding to the biotin moiety of the BIE substrate, increasing the effective local concentration of reactive iodoacetamide and thus raising the rate of nonspecific attack at neighbouring nucleophilic groups on the RNA. This mechanism would predict that (1) the original biotin binder (carrying the active site guanosine and possessing micromolar affinity for the biotin) would have significant alkylation activity, and (2) a selected ribozyme would have a low K_m to maximize substrate binding and local concentration effects. Experimental measurements, however, do not support either of these predictions. The biotin aptamer has no detectable self-alkylation activity, and kinetic measurements for the ribozyme-mediated reaction indicate a K_m for BIE of approximately 1 mM, substantially higher than the BIE concentration present during the ribozyme selection (200 μ M).

An alternative explanation for the alkylation activity would be that the RNA folds to generate an internal 'alkylation hot-spot', an exposed nucleotide accessible to free-floating electrophilic substrates. If this is the case, it should be possible to inactivate the ribozyme irreversibly with moderate concentra-

tions of competing iodoacetamide. Similarly, other electrophiles, such as dimethyl sulphate (DMS), should show a strong preference for reaction at the activated site. Neither of these predicted results are observed. Iodoacetamide (up to 10 mM, a 50-fold excess over BIE) has no effect on the BL2.8-7 biotinylation rate, and RNA incubated overnight with iodoacetamide shows no detectable reaction. The nucleophile for the ligation reaction (N7 of G⁹⁶) is no more reactive towards DMS (commonly used to probe the accessibility of guanosine N7 positions) than other guanines within the ribozyme (Fig. 3). Although the possibility remains that substrate binding induces conformational changes that dramatically increase the active site nucleophilicity, these results suggest that the basic model of a preformed active site ready to react with any available electrophilic substrate is essentially incorrect.

A third mechanism to account for catalytic activity involves very precise positioning within the ribozyme active site of the guanosine nucleophile and the iodoacetamide acceptor methylene. The rate enhancement due to positioning effects alone can be estimated by extrapolating from the background reaction rate to the catalysed rate. Kinetic measurements place k_{cat} for the BL2.8-7 enzyme, 0.0104 s⁻¹, at 2.3×10^7 times the

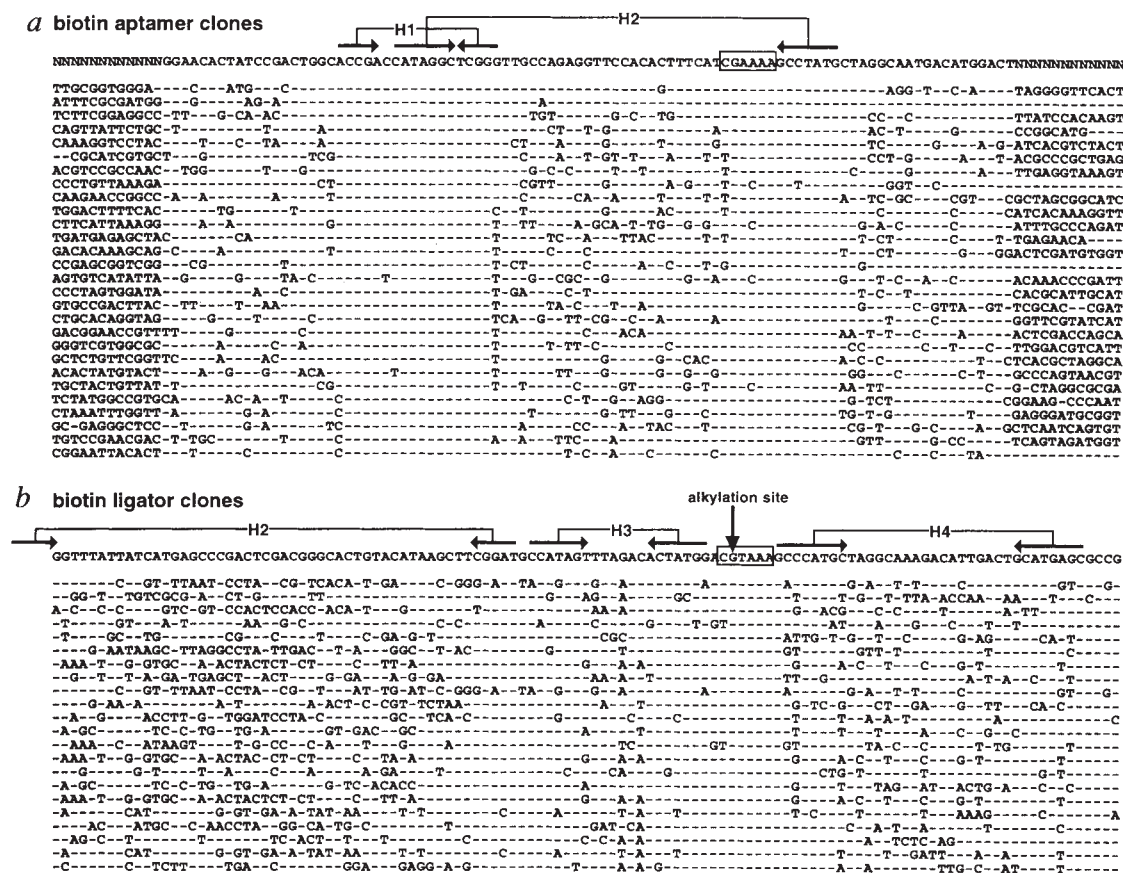


FIG. 4 Functional *a*, biotin binder, and *b*, biotin ligator sequences. Arrows indicate the locations of proposed helices. Boxed nucleotides are highly conserved yet not involved in secondary structure.

METHODS. PCR-amplified DNA was subcloned into the pCR vector using the TA cloning kit (Invitrogen, Inc.). Clones were sequenced by the Sanger dideoxynucleotide method using the universal M-13 primer. The partially randomized pool sequence is shown above each set of sequences. Deviations from the principle nucleotide at each position

are written explicitly and conservation of the wild-type base is indicated with a dash. Biotin aptamer and self-biotinylating RNA partially randomized pools were reselected for biotin-agarose binding and self-biotinylation, respectively. Biotin aptamer sequences correspond to clones from the fourth round of reselection. Self-biotinylating ribozyme clones were sequenced after eight rounds of reselection, when the overall biotinylation activity of the pool was 100 times the activity of the initial BL8-6 ribozyme.

non-catalysed rate, $4.6 \times 10^{-10} \text{ s}^{-1}$, measured at $200 \mu\text{M}$ BIE. Extrapolating from these values, an effective BIE concentration of approximately $4,500 \text{ M}$ in the ribozyme active site would be required if positioning effects alone are responsible for catalysis. This physically non-meaningful value (equivalent to confining the chemical reactants to a $0.7\text{-}\text{\AA}$ box) is nonetheless within the range of effective nucleophile concentrations observed for some intramolecular reactions in which translational degrees of freedom for the reactants have been restricted (often reaching 10^5 M)¹⁹. Positioning effects alone would seem sufficient to account for the observed catalytic activity, although a determination of their actual contribution to the total catalytic enhancement will require additional enzymological study.

Structural differences

Given that the biotin ligator arose by mutagenesis of the biotin binder sequence, and that both molecules interact specifically with biotin, we expected to find significant structural similarities between the two RNAs. Simple comparison of their primary sequences, however, failed to identify a well-conserved domain that might play a functional role; mutations appear randomly distributed along the length of the two sequences. To characterize the functional cores of the two molecules, we analysed the sequences of active clones isolated from the two mutagenized RNA pools generated from the biotin aptamer and self-alkylating ribozyme sequences. After four rounds of reselection with

the biotin aptamer-derived pool, more than 40% of the applied RNA bound tightly to biotin agarose. Similarly, three rounds of reselection of the self-alkylating ribozyme-derived pool yielded a collection of RNAs with activity matching that of the original BL8-6 clone, and five additional rounds of selection increased the activity ~100-fold. Approximately 30 individual RNAs from each of these subcloned pools were sequenced and analysed to determine which nucleotide positions were conserved and which pairs of nucleotides covaried to maintain Watson–Crick base pairing. The results of these experiments are summarized below and in Figs 4 and 5.

Two regions of the biotin binder are very highly conserved in clones that retain binding activity (Fig. 4a). Mutations at the 5'- and 3'-ends of the first conserved domain (changing the A⁵³·G⁷⁰ pair to either C·G or A·T) suggest a hairpin structure stabilized by a four-base-pair Watson-Crick duplex. Seven non-paired bases in the middle of the first domain directly complement the 3'-terminal half of the second conserved domain, thus suggesting a pseudoknot structure (Fig. 5a). In that the bases in these conserved domains are essentially invariant, the sequence data provide no covariational evidence for the pseudoknot. To test the proposed structure, a series of site-directed mutants was generated and assayed for binding to biotin agarose. Single-base substitutions that disrupt proposed Watson-Crick base pairs in the pseudoknot completely abolish biotin binding, but compensatory second-site mutations that introduce non-native Watson-Crick base pairs are largely able to restore

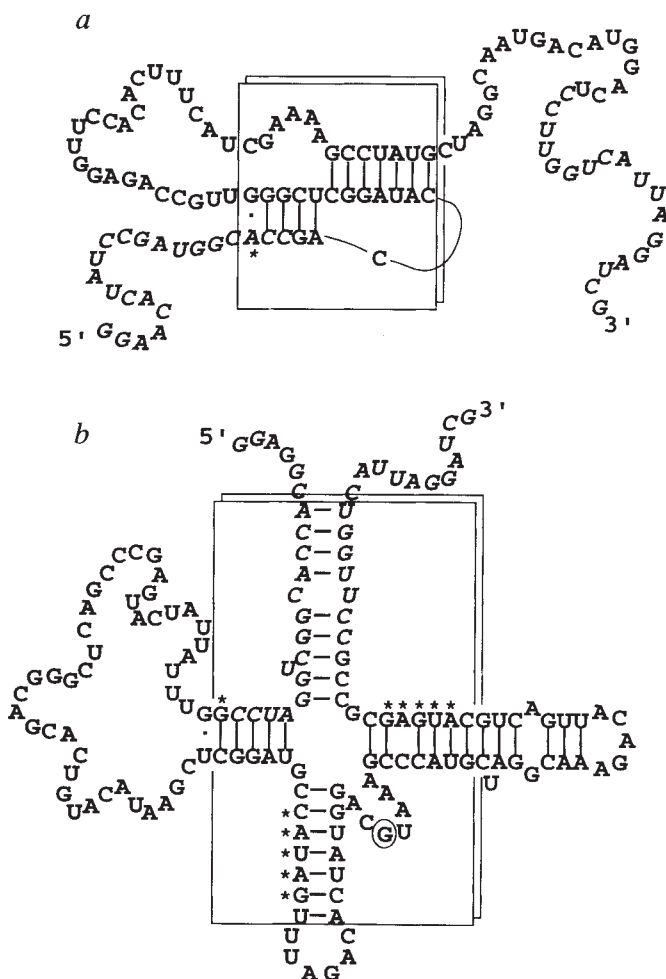


FIG. 5 Proposed secondary structures for the biotin aptamer and the self-biotinylating ribozyme. Nucleotides within the boxed region are highly conserved and make up the essential core of the aptamer and ribozyme. Asterisks indicate pairs of positions that covary in a Watson-Crick sense. Nucleotides in the constant primer sequences are shown in italics. *a*, Complete sequence of the BB8-5 biotin aptamer, shown as the proposed pseudoknot. *b*, Sequence and proposed cloverleaf structure for the BL8-6 self-biotinylating ribozyme. The guanosine residue that serves as the alkylation site for the biotinylation reaction is circled.

biotin binding. These data strongly support the proposed pseudoknot model for the biotin aptamer.

Comparison of the sequences of active ribozymes from the BL8-6 reselection indicate a striking change in structure relative to the original biotin binder. Nucleotides involved in the pseudoknot base pairing (53 to 70, 101 to 107), virtually invariant in the biotin binders, are poorly conserved in the enzyme sequences (Fig. 4*b*). In contrast, the ribozyme sequence in the region corresponding to the variable connecting loop of the biotin binder (nucleotides 71 to 94) appears to be well conserved, suggesting a structural role. Nucleotides that are very highly conserved in the biotin binder but not involved in the pseudoknot base pairing (... 5'-⁹⁵CGAAAAG¹⁰¹-3'...) are retained in the self-alkylating enzymes but with a highly conserved change to ... 5'-⁹⁵CGUAAAG¹⁰¹-3'... These results suggest that the change in function from biotin binding to alkylation of RNA with BIE is achieved by major structural rearrangements.

Further analysis of the BL8-6 derived sequences suggested a cloverleaf structure with several remarkable similarities to tRNA (Fig. 5*b*). The sequence ... 5'-⁹⁴ACGUAAA¹⁰⁰-3'... is presented as the tRNA variable loop, flanked on either side by extended duplexes (as indicated by several observed Watson-

Crick covariations). The single guanosine in the variable stem serves as the internal alkylation site for the enzyme. One interpretation of these results is that the hexanucleotide segments CGAAAA and CGUAAA directly mediate the interaction with biotin in the biotin binder and the biotin ligator, respectively, although they are presented in strikingly different secondary structure contexts. Comparison of ribozyme sequences from the third and eighth rounds of reselection suggest that the increase in pool alkylation activity is achieved by optimization of Watson-Crick base pairing in the cloverleaf duplexes and an increased fraction of purines (particularly adenosine) in the loop that caps helix 3 (analogous to the anticodon loop of tRNA).

To test the cloverleaf model for the biotin ligator, a synthetic ribozyme was designed by modifying one of the reselected sequences such that (1) primer sequences at the 5' and 3' ends not involved in the cloverleaf were deleted, (2) non-conserved bulges in the putative helices were removed, and (3) the variable loop of approximately 45 nucleotides (analogous to the DHU-loop) was replaced by a three-nucleotide loop sequence. This highly simplified ribozyme has approximately tenfold lower activity than the best reselected clone, but is still approximately tenfold more active than the original BL8-6 ribozyme, thus supporting the proposed cloverleaf structure. As a further test of this structure, we converted the simplified ribozyme into a two-piece system by effectively joining the 5' and 3' ends and breaking the loops at the ends of the helices that flank the alkylation site guanosine (analogous to the anticodon loop and the T Ψ C-loop). If the model for the ribozyme is correct, we would expect to find no catalytic activity for the individual fragments, but to be able to reconstitute activity by mixing the two fragments together (their association being driven by extensive base pairing). We find that the small fragment containing the active site guanosine does not self-biotinylate upon extended incubation with excess BIE. Addition of the large fragment, however, promotes stoichiometric biotinylation of the small fragment (data not shown). This experiment further confirms the proposed tRNA-like structure, and indicates that it is possible to engineer the ribozyme into a *trans*-acting enzyme.

We have shown that it is possible to evolve highly efficient ribozymes in a stepwise *in vitro* process starting from a pool of random sequence RNA. In contrast to previously characterized ribozymes, this class of *in vitro* selected enzymes is capable of binding a non-nucleic acid substrate (BIE) and performing a chemical reaction (*N*-alkylation) that is quite distinct from phosphate chemistry. This demonstration raises the possibility that RNA enzymes for a broad range of chemical reactions may be obtained by similar *in vitro* selection strategies, and supports the proposal that such ribozymes may have played an important role in the early evolution of life. □

Received 3 November 1994; accepted 6 March 1995.

- Joyce, G. F. *Nature* **338**, 217-224 (1989).
- Benner, S. A., Ellington, A. D. & Tauer, A. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7054-7058 (1989).
- Wilson, C. & Szostak, J. W. *Curr. Opin. struct. Biol.* **2**, 749-756 (1992).
- Ellington, A. D. & Szostak, J. W. *Nature* **348**, 818-822 (1990).
- Famulok, M. & Szostak, J. W. *J. Am. chem. Soc.* **114**, 3990-3991 (1992).
- Sassanfar, M. & Szostak, J. W. *Nature* **364**, 550-553 (1993).
- Connell, G. J., Illangsekere, M. & Yarus, M. *Biochemistry* **32**, 5497-5502 (1993).
- Lorsch, J. R. & Szostak, J. W. *Biochemistry* **33**, 973-982 (1994).
- Jenison, R. D., Gill, S. C., Pardi, A. & Polisky, B. *Science* **263**, 1425-1429 (1994).
- Famulok, M. *J. Am. chem. Soc.* **116**, 1698-1706 (1994).
- Bartel, D. P. & Szostak, J. W. *Science* **261**, 1411-1418 (1993).
- Lorsch, J. R. & Szostak, J. W. *Nature* **371**, 31-36 (1994).
- Prudent, J. R., Uno, T. & Schultz, P. G. *Science* **264**, 1924-1927 (1994).
- Green, N. M. *Adv. Protein Chem.* **29**, 85-133 (1975).
- Al-hakim, A. H. *Biochem. J.* **251**, 935-938 (1988).
- Peattie, D. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1760-1764 (1979).
- Tramontano, A., Ammann, A. A. & Lerner, R. A. *J. Am. chem. Soc.* **110**, 2282-2283 (1988).
- Janda, K. D., Weinhouse, M. I., Schloeder, D. M., Lerner, R. A. & Benkovic, S. J. *J. Am. chem. Soc.* **112**, 1274-1275 (1990).
- Fersht, A. *Enzyme Structure and Mechanism*, 2nd edn 57-59 (Freeman, New York, 1985).

ACKNOWLEDGEMENTS. This work was supported by grants from Hoechst A. G. and the NIH. C.W. was supported by the Jane Coffin Childs Memorial Fund for Medical Research. We thank our colleagues, in particular J. Lorsch and D. Bartel, for discussion and advice.

Discovery and physical properties of Dactyl, a satellite of asteroid 243 Ida

C. R. Chapman*, J. Veverka†, P. C. Thomas†, K. Klaasen‡, M. J. S. Belton§, A. Harch¶, A. McEwen¶, T. V. Johnson‡, P. Helfenstein†, M. E. Davies*, W. J. Merline* & T. Denk*

* Planetary Science Institute/SJI, 620 N. 6th Avenue, Tucson, Arizona 85705, USA

† Center for Radiophysics and Space Research, Cornell University, Ithaca, New York 14853, USA

‡ Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA

§ National Optical Astronomy Observatories, Tucson, Arizona 85719, USA

¶ US Geological Survey, Flagstaff, Arizona 86001, USA

* RAND, Santa Monica, California 90406, USA

* Institut für Planetenerkundung, 12484 Berlin, Germany

OBSERVATIONS of stellar occultations by asteroids have suggested that some may have satellites¹. But given the absence of any confirmatory evidence, the prevailing view has been that although such satellites probably do exist, they are likely to be rare². Here we report the discovery³ by the Galileo spacecraft of a satellite associated with the asteroid 243 Ida. Although the satellite, Dactyl, is only 1.6 km across, it has been imaged with sufficient resolution for geological analysis. We describe the physical properties of Dactyl, with emphasis on its notably smooth shape, its crater population (which includes a crater chain) and its photometric properties. We find that, spectroscopically, Dactyl resembles both Ida and the other members of the Koronis asteroid family, implying a similar composition; small spectral differences may reflect a space weathering process that slightly alters the colours with time. We argue that Dactyl originated during the breakup of the Koronis parent body, and that satellites could be common around other asteroids (particularly members of asteroid families).

Dactyl was discovered in one image of the final six-colour sequence of Ida images, taken at 16:38 UTC on 28 August 1993 through a green filter from a distance of 10,927 km. Dactyl is seen in 46 additional images, taken at 18 different times over a period of 5.39 h. All have been measured for position and brightness (and colour from colour sequences). One image, part of the high-resolution Ida mosaic⁴, has a resolution of 39 m per pixel at 48° phase and provides the most detailed view for analysis of Dactyl's surface geology (Fig. 1a). An even higher resolution picture, taken near encounter, primarily shows Dactyl's shadowed side (Fig. 1b) which is sufficiently illuminated by Ida ('Ida shine') to define its profile.

The remaining Ida imaging data have been examined for additional satellites. We searched much of the inner portion of Ida's gravitational sphere of influence, particularly along Ida's equatorial plane. The chief uncertainties are confusion by sporadic cosmic-ray hits on the CCD detector, requiring satellites to be revealed by motion in successive frames. Several interesting features appear to be anomalous cosmic-ray hits, but no other satellite has yet been found. It is unlikely that we have missed another object approaching Dactyl's size in the volume searched. (The detection threshold diameter within 10,000 km of Ida is ~50 m.)

Dactyl's size and shape have been determined using the Spud shape-modelling program⁵. Dactyl is approximated by an ellipsoid with axes of $1.6(\pm 0.1) \times 1.4(\pm 0.14) \times 1.2(\pm 0.24)$ km, or a total volume of $1.4(\pm 0.5)$ km³. The longest axis pointed towards

Ida (to $\pm 15^\circ$) at the time it was spatially resolved, and the shortest axis was perpendicular to its orbital plane.

Dactyl was observed for a duration that was greater than a complete Ida rotation but which was only a modest fraction of even the shortest of Dactyl's allowed orbital periods (~25 h)⁶. Its brightness hardly varied. Nor is there any other evidence for rotation, so we conclude that Dactyl's spin period is ≥ 8 h, consistent with the dynamical expectation that Dactyl should have decayed rapidly to tidal lock.

Dactyl does not look like an unmodified, angular rock fragment (Fig. 1a, b). Despite a slight concavity indicated by the trend of its terminator in Fig. 1a, Dactyl has a rather smooth, regular shape, which contrasts with the angularity of Ida and Gaspra and many other small bodies (for example, comet Halley, and the asteroids Toutatis and Castalia). On the other hand, the martian moons Phobos and Deimos present some aspects with comparable (even smaller) limb roughnesses⁷, so Dactyl's regular shape is not wholly unprecedented. Local parts of the surfaces of Ida and Gaspra, comparable in scale to Dactyl, may be similarly smooth.

Images of Dactyl over phase angles 20°–96° permit determinations of its albedo and photometric properties. Within error bars, Dactyl's geometric albedo is identical to Ida's (0.20 ± 0.02 at $0.56 \mu\text{m}$) as is its phase function (it brightens 0.028 ± 0.002 mag per degree of decreasing phase angle), consistent with Dactyl's surface having a texture similar to average Ida. Therefore the data seem to indicate that Dactyl's surface is not regolith-free bare rock.

The visible and near-infrared reflectance spectra of Dactyl and Ida are similar overall (Fig. 2), implying that they have similar compositions (Ida itself is in the S(IV) subgroup defined by Gaffey⁸). This is consistent with the hypothesis that Dactyl, as well as Ida, originally derived from the parent body of the Koronis asteroid family. Koronis family members are known to have similar colours⁹, implying rather homogeneous composition of the parent body.

Dactyl shows real minor spectral differences from Ida, however, including its slightly less-red colour. The differences are of similar magnitude to variations among various Koronis family members¹⁰. However, Dactyl's spectrum tends to differ from other well observed Koronis members in an intriguing way. Dactyl has a significantly deeper 1- μm absorption band than Ida (and nearly all other Koronis members) and it shows slightly greater spectral curvature than all observed Koronis members (that is, its relative reflectance is lower in the violet and infrared than at intermediate wavelengths). These differences, which should be verified by further spectrophotometry of Koronis members, are generally in the same sense that smaller S-type asteroids at smaller heliocentric distances differ from their larger, more distant siblings^{8,11}.

The spectral differences between Dactyl and Ida (and other Koronis members) could arise from slight compositional differences and/or from a 'space weathering' process that slightly modifies surface colours with time. (Another explanation, differences in particle sizes, seems unlikely given the identical photometric properties of Ida and Dactyl.) Real compositional differences would mean either that Dactyl could not have been derived directly from Ida or that there was heterogeneity in the Koronis parent body, whereas a weathering process could make identical material look different if weathered for different lengths of time or at different rates.

A space weathering process, like that operating on the Moon but much weaker¹², which 'reddens' and linearizes spectra and weakens absorption bands, would convert a Dactyl-like spectrum to be Ida-like. That such a process actually operates on Ida (even though we do not know how it works) is indicated by the fact that fresh, young craters on Ida (called terrain B) are spectrally intermediate between Dactyl and most of Ida (terrain A)^{13,14}. Thus Dactyl could represent an intermediate stage in weathering between an initial mineral assemblage representative

¶ Present address: Center for Radiophysics and Space Research, Cornell University, Ithaca, New York 14853, USA.

of the Koronis family and progressing to a 'mature' Ida. Small S-type asteroids, like Dactyl, may be less weathered because they are younger or because their surface layers are more frequently jolted and rearranged by impacts. Alternatively, they may be less subject to a weathering process dependent on gravity (for example, regolith maturation).

The most notable geological features of Dactyl are craters. We have identified up to 29 possible craters, of which 15 are definite. From a region of favourable lighting and visibility (basically the half of the image in Fig. 1a closest to the terminator), we have determined the cumulative diameter-frequency relation for the definite craters—it is consistent (in both slope and overall crater density) with the small end of Ida's crater population, which reflects equilibrium with saturation cratering⁴.

Dactyl's craters may not be simply due to impacts by heliocentric projectiles. At least three sizeable craters (and up to seven, including 'possible' craters) lie in a well defined crater chain. It could have been produced by an aligned jet of second-

ary ejecta, from a larger cratering event on Ida, which immediately struck Dactyl. Although some crater chains on larger bodies are of internal origin, this seems unlikely for a body as small as Dactyl.

There is a positive relief feature within the most prominent, large crater. (We avoid the term 'central peak' because such peaks form only in much larger craters.) Possibly, the feature is the wall of a smaller crater nestled within the larger crater. The most intriguing interpretation is that the feature is the remnant of the projectile that formed the crater, if it struck at very low velocity. The ratio of projectile-to-crater size is consistent with sub-escape velocities within the Ida system.

We see no evidence of grooves, ridges, sharp edges, or other geological features in our best images.

It is highly improbable that Dactyl was captured from an independent heliocentric orbit by Ida, because there is no plausible source of dissipation (collision with a pre-existing satellite begs the question of how *that* satellite could have formed). A

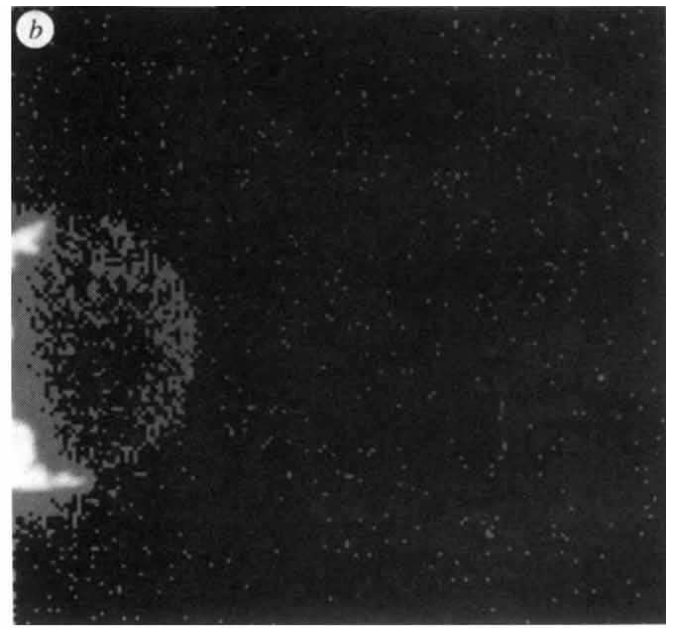
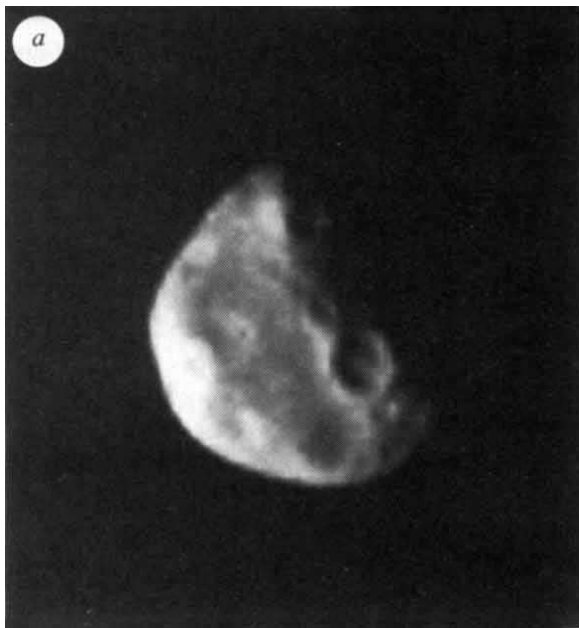


FIG. 1 a, High-resolution image of Dactyl, showing craters and other surface topography discussed in the text. b, Dactyl image on the edge of one frame from the partial mosaic taken near the time of Ida encounter.

Except for the horns of Dactyl's crescent, the Sun-illuminated part of Dactyl is beyond the edge of the frame. The shadowed side of Dactyl can be seen illuminated by 'Ida-shine'.

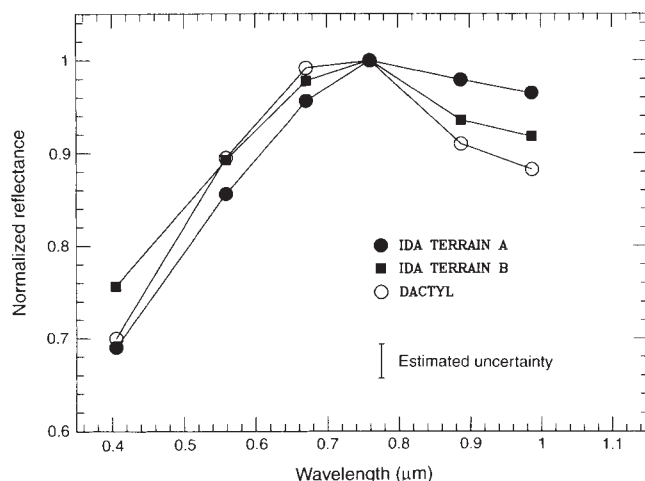


FIG. 2 Comparison of the spectral reflectances of Dactyl and Ida, measured from the final six-colour imaging sequence. Two major colour units (terrains A and B) have been identified on Ida. Much of the face of Ida seen in the higher-resolution views is dominated by terrain A. The data are normalized to unity at a wavelength of 0.756 μm. The typical uncertainty ($\pm 1.9\%$) of these relative colours represents the accuracy to which the colour ratio of a single feature on Ida can be duplicated from one image sequence to another, determined from a series of features tracked in several Ida colour sequences.

primordial origin for asteroidal satellites has been suggested¹⁵, but an object as small as Dactyl is very unlikely to survive over the age of the Solar System. Almost certainly Dactyl was formed either from Ida (for example, from re-accretion of ejecta from a large crater) or else at the same time as Ida during the catastrophic collision that formed the Koronis family. Both ideas are compatible with the compositional similarity between Ida and Dactyl.

There are intriguing similarities between the properties of Dactyl and those predicted² for a satellite formed from cratering ejecta (a small object in a prograde orbit around a rapidly spinning asteroid). However, it was also concluded² that 'satellites formed by this mechanism may be rare or nonexistent' and large harmonics in Ida's gravitational potential due to its irregular shape further mitigate against orbital stability of an object formed so close to Ida.

The formation of binary asteroids or asteroidal satellites during a catastrophic collision was originally suggested by Hartmann¹⁶ and proposed more specifically as an outcome of natural 'jetting' phenomena studied experimentally by Martelli *et al.*¹⁷. We prefer this explanation for the origin of Dactyl. Although Durda¹⁸ has found a low rate of formation of binaries (where the two bodies are approximately the same size) by this mechanism, small satellites are more likely to be formed this way. Because Ida itself is saturated with craters⁴, and because of the shorter collisional lifetimes for smaller bodies, it is somewhat unlikely that Dactyl could have avoided collisional disruption since its formation. It is more likely to be a remnant from a pre-existing, perhaps larger, original satellite that has undergone breakup and re-accumulation from orbiting debris. Therefore, Dactyl could well have a 'rubble-pile' structure even if it formed originally as a solid Koronis fragment.

Dactyl's relatively smooth shape could reflect several generations of breakup and re-accumulation, perhaps augmented by the sweeping up of some other temporary debris ejected from Ida. However, if Dactyl managed to survive breakup, it could well have been smoothed by erosion. The steeply sloping power-law size distribution of asteroidal projectiles found from Gaspra's crater population¹⁹, if it extends to sufficiently small sizes, would result in a 'sandblasting' erosional smoothing. Whether Dactyl is solid or composed of rubble, its smooth shape (and Ida's smoothness at comparable scales) probably reflects, in part, the apparently steep size distribution for heliocentric cratering projectiles ~10–100 m in diameter (that is, there is a relative lack of projectiles capable of making large gouges in Dactyl).

As we noted, some of Dactyl's craters may be locally derived and one may be from a low-velocity impact. If true, the assumption of Belton *et al.*⁴ that Ida has been struck solely by heliocentric projectiles from the same population that cratered Gaspra may be too simplistic. If cratering projectiles are derived from within the Ida system, or if the Koronis precursor breakup produced a high flux of Koronis-derived projectiles, then Gaspra-based calculations of Ida's age may be upper limits, and Ida could be younger, more in accord with other age estimates for the Koronis family^{20,21}.

The discovery of Dactyl (which would be difficult to detect from Earth), together with other data (indications from radar that some small asteroids may be contact binaries, studies of double craters, and so on^{22,23}), has caused reassessment of how common asteroidal satellites may be. Unconfirmed photometric blink-outs associated with asteroidal stellar occultations remain inadequate proof of the existence of any particular satellites. Nevertheless, Dactyl's existence suggests that asteroidal satellites may be more common (especially for family members) than had been thought. Discovery of additional asteroidal satellites would provide important clues about the nature of asteroids (for example, the mass of the primary asteroid) and about asteroidal processes. □

Received 1 February; accepted 31 March 1995.

1. Van Flandern, T. C., Tedesco, E. F. & Binzel, R. P. in *Asteroids* (ed. Gehrels, T.) 443–465 (Univ. Ariz. Press, Tucson, 1979).
2. Weidenschilling, S. J., Paolicchi, P. & Zappalà, V. in *Asteroids II* (eds Gehrels, T., Binzel, R. P. & Matthews, M. S.) 643–658 (Univ. Ariz. Press, Tucson, 1989).
3. Belton, M. J. S. & Carlson, R. *IAU Circ. No.* 5948 (1994).
4. Belton, M. J. S. *et al. Science* **265**, 1543–1547 (1994).
5. Simonelli, D. P., Thomas, P. C., Carcich, B. T. & Veverka, J. *Icarus* **103**, 49–61.
6. Belton, M. J. S. *et al. Nature* **374**, 785–788 (1995).
7. Thomas, P. *Icarus* **77**, 248–274 (1989).
8. Gaffey, M. J. *et al. Icarus* **106**, 573–602 (1993).
9. Gradie, J. C., Chapman, C. R. & Williams, J. G. in *Asteroids* (ed. Gehrels, T.) 359–390 (Univ. Ariz. Press, Tucson, 1979).
10. Binzel, R. P., Xu, S. & Bus, S. J. *Icarus* **106**, 608–611 (1993).
11. Dermott, S. F., Gradie, J. & Murray, C. D. *Icarus* **62**, 289–297 (1985).
12. Pieters, C. M., Fischer, E. M., Rode, O. & Basu, A. *J. geophys. Res.* **98**, 20817–20824 (1993).
13. Veverka, J. *et al. Icarus* (submitted).
14. Sullivan, R. *et al. Icarus* (submitted).
15. Gehrels, T., Drummond, J. & Levenson, N. *Icarus* **70**, 257–263 (1987).
16. Hartmann, W. K. in *Asteroids* (ed. Gehrels, T.) 466–479 (Univ. Ariz. Press, Tucson, 1979).
17. Martelli, G. *et al. Astr. Astrophys.* **271**, 315–318 (1993).
18. Durda, D. D. *Bull. Am. astr. Soc.* **26**, 1158 (1994).
19. Belton, M. J. S. *et al. Science* **257**, 1647–1652 (1992).
20. Binzel, R. P. *Icarus* **73**, 303–313 (1988).
21. Marzari, F., Davis, D. & Vanzani, V. *Icarus* **113**, 168–187 (1995).
22. Ostro, S. J. *et al. Science* **248**, 1523–1528 (1990).
23. Bottke, W. F. & Melosh, H. J. *Lunar planet. Sci.* **XXVI**, 153–154 (1995).

ACKNOWLEDGEMENTS. We thank R. Binzel and M. Gaffey, colleagues including S. Weidenschilling and J. Gradie, and other members of the Galileo Project for their help. The Galileo Project is managed by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA.

Bulk density of asteroid 243 Ida from the orbit of its satellite Dactyl

M. J. S. Belton*, **C. R. Chapman†**, **P. C. Thomas‡**,
M. E. Davies§, **R. Greenberg||**, **K. Klaassen¶**,
D. Byrnes¶, **L. D'Amario¶**, **S. Synnott¶**,
T. V. Johnson¶, **A. McEwen***, **W. J. Merline†**,
D. R. Davis†, **J.-M. Petit||**, **A. Storrs****, **J. Veverka‡**
& **B. Zellner††**

* National Optical Astronomy Observatories, Tucson, Arizona 85719, USA

† Planetary Science Institute/SAIC, Tucson, Arizona 85705, USA

‡ Center for Radiophysics and Space Research, Cornell University, Ithaca, New York 14853, USA

§ RAND, Santa Monica, California 90406, USA

|| Lunar and Planetary Laboratory, University of Arizona, Tucson, Arizona 87721, USA

¶ Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA

* United States Geological Survey, Flagstaff, Arizona 86001, USA

** Space Telescope Science Institute, Baltimore, Maryland 21218, USA

†† Department of Physics, Georgia Southern University, Statesboro, Georgia 30460, USA

DURING its reconnaissance of the asteroid 243 Ida, the Galileo spacecraft returned images of a second object, 1993(243)1 Dactyl¹—the first confirmed satellite of an asteroid. Sufficient data were obtained on the motion of Dactyl to determine its orbit as a function of Ida's mass. Here we apply statistical and dynamical arguments to constrain the range of possible orbits, and hence the mass of Ida. Combined with the volume of Ida², this yields a bulk density of $2.6 \pm 0.5 \text{ g cm}^{-3}$. Allowing for the uncertainty in the porosity of Ida, this density range is consistent with a bulk chondritic composition, and argues against some (but not all) classes of meteoritic igneous rock types that have been suggested as compositionally representative of S-type asteroids like Ida.

Figure 1 shows a selection from 47 independent views of Ida and Dactyl obtained by the Galileo solid-state imaging (SSI)

TABLE 1 Predicted bulk porosities for Ida

Meteorite type	Typical sample density (g cm ⁻³)	Typical sample porosity	Bulk porosity range implied for Ida
Iron	~7.9	~0	0.60–0.73
Stony-iron	~5.1	<0.01	0.40–0.59
Ordinary chondrite	~3.6	~0.11	0.23–0.48

Typical density and porosity of samples of major meteorite types²³ and the nominal bulk porosity implied for Ida if the asteroid is composed of one meteorite type.

camera³ that were spaced over a period of 5 h 24 min. It was quickly evident from the data that Dactyl was ~85 km from Ida, in the foreground of the images, moving slowly (~6 m s⁻¹) in the same direction as Ida's rotation (that is, prograde motion with respect to Ida's spin but retrograde motion with respect to Ida's orbital motion), in an orbit near Ida's equatorial plane.

We present an assessment of orbits that fit these data under the assumption of simple two-body motion. This assumption is justified for a preliminary calculation because the observed orbital arc was short (45–90°) and tests, using a spinning ellipsoidal potential emulating Ida, show maximum spatial displacements from two-body motion less than the diameter of Dactyl (~1.4 km) at the beginning of the arc when Dactyl was not spatially resolved. Dactyl's small volume, ~10⁻⁴ times that of Ida, ensures that the system centre of mass can be assumed to coincide with that of Ida.

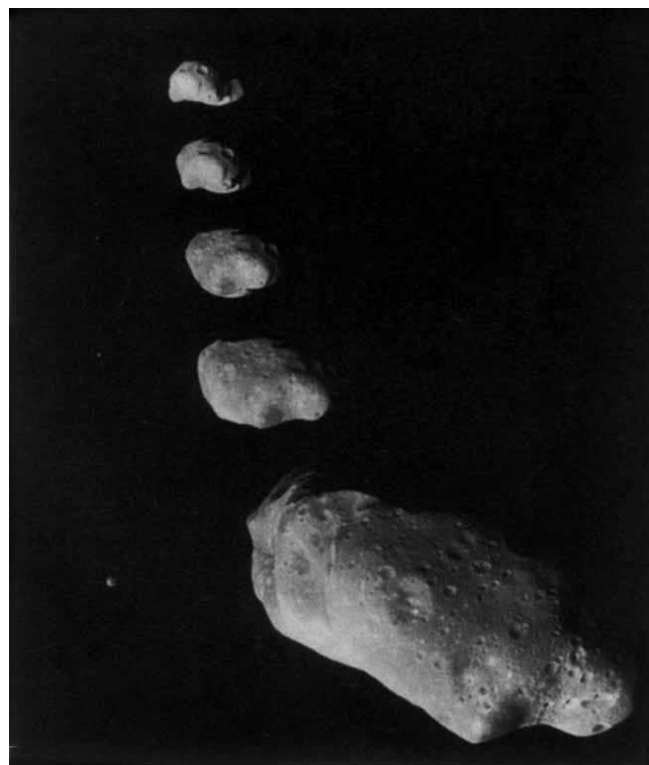


FIG. 1 A montage of views of the Ida/Dactyl system taken on 28 August 1993 with the solid-state imaging camera on the Galileo spacecraft. The sequence is arranged chronologically, starting with the most distant images at the top. The effects of decreasing range, rotation and increasing spatial resolution can easily be seen. In order, the timing and range to Ida of the images relative to closest approach is (h, km): -1.1, 50,406; -0.95, 42,148; -0.73, 33,140; -0.55, 24,918; -0.24, 10,927. (Dactyl is hard to see in the views at the top of the figure; but note that the positions of Dactyl are all in a line in this montage.)

Ida and Dactyl are spatially resolved in most images. Accurate knowledge of their shapes and orientations was therefore used for precise estimation of the positions of their projected centres of mass. Ida's interior is assumed to be homogeneous. Resolved images of Ida over a full rotation period (4.633 h)⁴ allowed detailed shape models² to be developed, yielding volumes of $16,100 \pm 1,900$ km³ for Ida (mean radius, 15.7 km) and 1.4 km³ for Dactyl (mean radius, 0.7 km). Ida's rotation pole was determined from SSI data⁵ to be in the direction right ascension 348.76°, declination 87.10° (J2000).

Figure 2 shows two-body orbits that fit the SSI data. Predicted positions of Dactyl relative to Ida were directly compared with measured positions in the image plane. Weighted root-mean-square (w.r.m.s.) errors of fit to the best-fitting orbit solutions were consistently <0.2 pixels in image coordinates. In Fig. 3 selected orbital elements, periape distance (R_p), and w.r.m.s. errors are shown as a function of GM , where G is the gravitational constant and M the mass of Ida. (Periape is the point in the orbit which is closest to the centre of mass.)

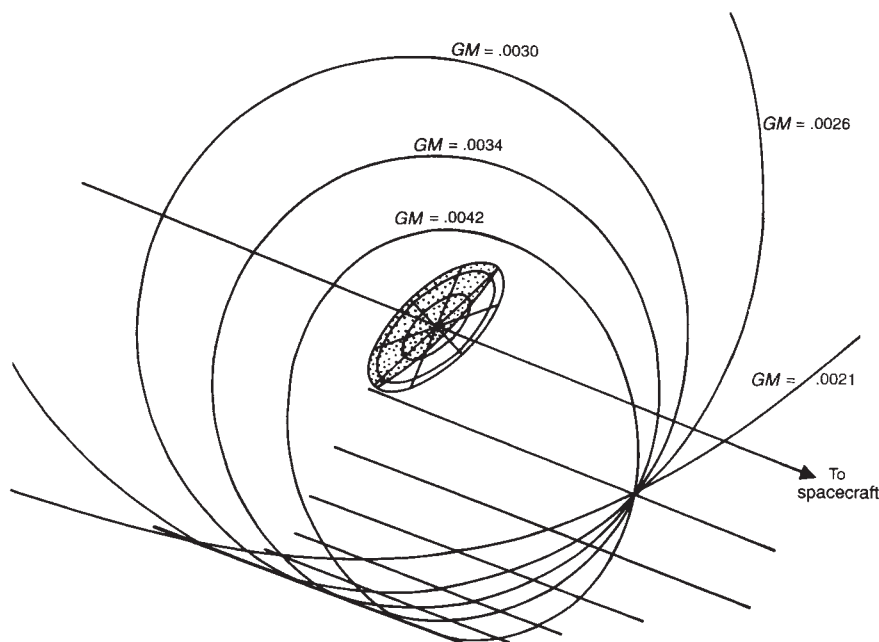
The w.r.m.s. errors show a slight preference for lower values of GM , but do not allow a specific value of GM to be determined with confidence. For $GM < 0.0023$ km³ s⁻², the orbits are highly elliptical (eccentricity $e > 0.98$) or hyperbolic with R_p decreasing from 78 to 70 km. We estimate the probability of the chance presence of any detectable (radius > 100 m) main-belt asteroid in Ida's entire gravitational sphere of influence (radius ~7,000 km) at only ~10⁻⁷. It is therefore highly unlikely that Dactyl is on a hyperbolic orbit and that GM is less than 0.0023 km³ s⁻².

Beyond a distance $R_T \approx a_D (M/M_\odot)^{2/5}$, where $M_\odot$ is the mass of the Sun and a_D is the heliocentric distance of the Dactyl–Ida system (that is, ~1,500 km), two-body motion is no longer a good approximation⁶. Hamilton and Burns⁷ indicate that circular retrograde (with respect to the orbital motion) orbits with semimajor axes as large as ~445 times the primary's mean radius (~7,000 km) may be long-lived, but the orbits for Dactyl are highly elliptical with their periape distance approaching Ida as their apoapse increases beyond R_T . (Apoapse is the point in the orbit which is farthest from the centre of mass.) This type of orbit should be sensitive to small perturbations and may possibly be chaotic, although regular orbits of this class may exist, locked-in by resonance effects⁸.

As Dactyl could be on a bound orbit that takes it relatively far from Ida, we searched for it in images (from the Wide Field Planetary Camera on the Hubble Space Telescope) that were taken on 1994 April 26.2 UT, ~240.5 d after the Galileo encounter. We found no trace of the satellite between projected distances of 400 and 30,000 km from Ida. Dactyl's estimated visual magnitude was 21.6, ~1 mag brighter than the detection limit of our data. A search to lower distances was prevented by scattered light from Ida. This excludes hyperbolic orbits for $0.0020 < GM < 0.0023$ km³ s⁻² and makes orbits with apoapses > 700 km (that is, $GM < 0.0024$ km³ s⁻²) unlikely.

Orbits with $GM > 0.0030$ km³ s⁻² (Fig. 3) become increasingly elliptical, with R_p rapidly falling towards distances comparable with the size³ of Ida ($a = 28$ km, $b = 12$ km, $c = 10.5$ km, where a, b, c are the principal axes of the best-fit ellipsoid). Calculations of the evolution of orbits about a triaxial ellipsoid (used to simulate Ida) show that orbits with $GM > 0.0031$ km³ s⁻² (that have $R_p < 75$ km) either collide with Ida or are ejected on a hyperbolic orbit. This result is consistent with the work of Chauvineau *et al.*⁸ and Scheeres^{9,10}, who showed that circular prograde orbits could be stable outside a critical distance from the primary (~63 km). For Dactyl to have a stable elliptical prograde orbit with $R_p < 75$ km it would have to be trapped in some special, low-order, resonance with Ida's spin period. In view of the distinctly non-ellipsoidal shape of Ida and the comments of Chauvineau *et al.*⁸ on the effects of deviations from an ellipsoidal shape, we doubt that such special orbits are appropriate for Dactyl. We conclude that, for the range of orbits that satisfy

FIG. 2 Possible orbits of Dactyl. Each orbit is labelled with the value of the assumed gravitational parameter, GM ($\text{km}^3 \text{s}^{-2}$). The outline of Ida is approximated by an ellipsoid and is drawn to scale. The orientation shown for Ida is for the time of closest approach, 1993 August 28.70284 yr . The view directions for several of the images are shown as diagonal lines. Because these directions lie close to both Dactyl's orbital plane and Ida's equatorial plane, and because there was little parallactic motion between Ida and Dactyl during the early parts of the sequence, it was not possible to distinguish between the best orbits for different GM . The orbits approximately intersect at $\sim 85 \text{ km}$ from the centre of Ida because the final images of the fly-by sequence were taken from a wide range of viewing angles but within a few minutes of each other, thus rendering Dactyl effectively stationary from different view points. These latter data therefore allow accurate triangulation of the position of Dactyl relative to Ida. The motion is prograde with respect to Ida's spin which is retrograde, that is, the view is from the direction of Ida's south pole.



the observations and are likely to be stable, GM is in the range $0.0024\text{--}0.0031 \text{ km}^3 \text{s}^{-2}$, that is, $GM = 0.0028 \pm 0.0004 \text{ km}^3 \text{s}^{-2}$. The corresponding mass of Ida is $M = 4.2 \pm 0.6 \times 10^{19} \text{ g}$, and the range of possible Dactyl orbital periods is 227.3–24.7 h. Radio tracking of the spacecraft during the encounter might have yielded an independent estimate of Ida's mass. However, no systematic Doppler shifts in the radio signal were detected (J. D. Anderson, personal communication).

Ida's bulk density, $\rho_B = 2.6 \pm 0.5 \text{ g cm}^{-3}$, follows immediately from the above estimates and gives insight into its bulk composition and porosity, p . These quantities are related to the grain density, ρ , of the rock type by $p = (1 - \rho_B/\rho)$.

S-type asteroids have been proposed^{11–13} as the parent bodies for ordinary chondrites, the most numerous meteorite type¹⁴. Although there are spectral similarities between the two, there are also strong differences¹⁵. In the absence of any demonstrated 'space weathering' process to transform suitably the spectrum of ordinary chondritic material^{16–18}, the current assessment is that most S-type asteroids are closely related to types of stony-iron meteorites that reflect a record of igneous processes in the primitive planetary system. This identification leaves the relatively primitive and abundant material in ordinary chondrites without any well observed parent-body type—a situation sometimes referred to¹⁹ as the "ordinary-chondrite mystery". Meteoritic analogues to the S-types that have been suggested^{19–22} include pallasites, olivine-dominated stony iron, lodranites, irons, and, for a small subset of S-types, ordinary chondritic material. In Table 1 we give estimates of Ida's bulk porosity using typical densities and porosities of meteoritic types²³ assuming that they dominate Ida's composition.

Bulk porosities for Ida, which may be appreciably fractured in its interior, are expected to be larger than the porosities seen in meteorite samples. Thus an Ida of stony-iron composition (typically $p < 0.01$) is expected to have a bulk porosity of $>1\%$, whereas a chondritic composition (typically $p \approx 0.11$) is expected to have a bulk porosity $>11\%$. Estimates of the bulk density of other Solar System objects of comparable size to Ida, for example, Janus, Empimetheus, Phobos and Deimos, when interpreted in terms of reasonable assumed bulk compositions, suggest that bulk porosities in the range 0.1–0.6 are plausible (we

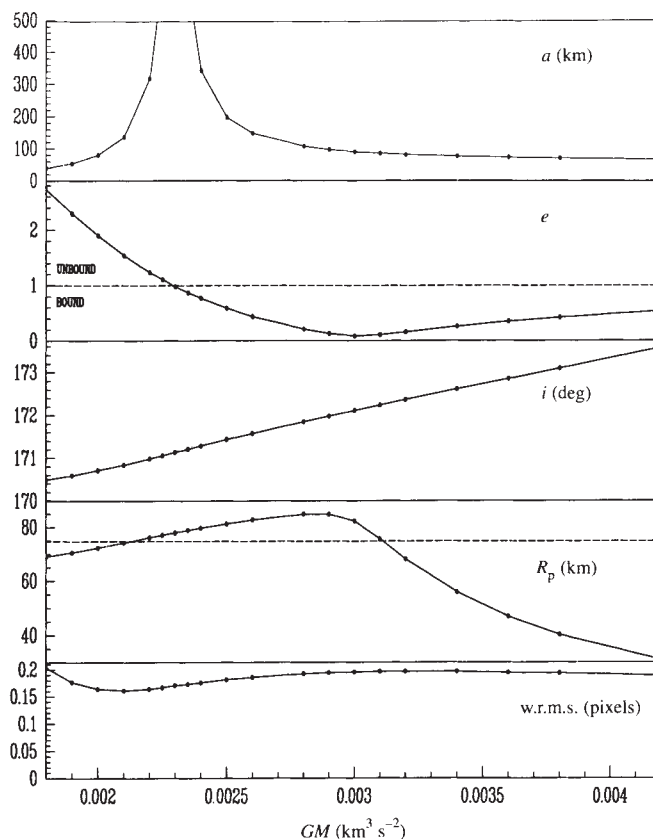


FIG. 3 Variation of orbit shape, inclination, and goodness of fit with assumed value of Ida's gravitational parameter, GM . Plotted variables (top panel to bottom panel) are as follows. Semimajor axis, a ; for $GM < 0.0023 \text{ km}^3/\text{sec}^2$, the orbits are hyperbolic. Eccentricity, e . Inclination (with respect to Ida's equatorial system), i . Periape distance, R_p —bound orbits with R_p below the dashed line at 75 km are found to collide with, or escape from, Ida on very short timescales. Weighted root-mean-square, w.r.m.s., sum of deviations of orbital positions from observed locations.

consider the upper limit as an extreme driven by the uncertain density of the smallest of these objects, Deimos). A comparison with known conditions typical of the lunar surface is also instructive: the overburden pressure in the interior of Ida should be similar to that at ~ 100 m depth on the Moon, that is, well below the fluffy outer layers for which we have measurements²⁴ (for example, $p \approx 0.44$ for depths of 40–60 cm) and in a region of increasing compaction with depth. Neglecting compressive effects on the particulates themselves, such soils may compact to interparticle porosities approaching a theoretical limit of 0.26. We therefore suggest that bulk porosities >0.40 are unlikely for an object as large as Ida even if heavily fractured. Accepting this constraint, the results in Table 1 then point away from sub-classes of stony-irons (that is, certain iron-rich pallasitic and other heavy igneous analogues for Ida's bulk composition) and a stony bulk composition with moderate NiFe content appears to be favoured. However, only a pure iron composition seems to be definitely ruled out. $\square$

Received 11 January 1995; accepted 3 March 1995.

1. Belton, M. J. S. & Carlson, R. *IAU Circ. No.* 5948 (1994).
2. Thomas, P. et al. (abstr.) *Bull. Am. Astr. Soc.* **26**, 1155 (1994).
3. Belton, M. J. S. et al. *Science* **265**, 1543–1547 (1994).
4. Binzel, R. P. et al. *Icarus* **105**, 310–325 (1993).

5. Davies, M. E. et al. (abstr.) *Bull. Am. Astr. Soc.* **26**, 1154–1155 (1994).
6. Roy, A. E. *Orbital Motion* 3rd edn 170 (Hilger, Bristol, 1988).
7. Hamilton, D. P. & Burns, J. A. *Icarus* **96**, 43–64 (1992).
8. Chauvineau, B., Farinella, P. & Mignard, F. *Icarus* **105**, 370–384 (1993).
9. Scheeres, D. J. *Icarus* **110**, 225–238 (1994).
10. Scheeres, D. J., Ostro, S. J., Hudson, R. S. & Werner, R. A. (abstr.) *Bull. Am. Astr. Soc.* **26**, 1166–1167 (1994).
11. McCord, T. B. & Chapman, C. R. *Astrophys. J.* **195**, 553–562 (1975).
12. Leake, M., Gradie, J. & Morrison, D. *Meteoritics* **13**, 101–120 (1978).
13. Feierberg, M. A., Larson, H. P. & Chapman, C. R. *Astrophys. J.* **257**, 361–372 (1982).
14. Sears, D. W. G. & Dodd, R. T. in *Meteorites and the Early Solar System* (eds Kerridge, J. F. & Mathews, M. S.) 3–31 (Univ. Arizona Press, Tucson, 1988).
15. Gaffey, M. J. *Icarus* **60**, 83–114 (1984).
16. Clarke, B. E., Fanale, F. P. & Salisbury, J. W. *Icarus* **97**, 288–297 (1992).
17. Fanale, F. P., Clark, B. E. & Bell, J. F. *J. geophys. Res.* **97**, 20863–20874 (1992).
18. Bell, J. F. & Keil, K. *Proc. lunar planet. Sci. Conf.* **18**, 573–580 (1988).
19. Bell, J. F., Davis, D. R., Hartmann, W. K. & Gaffey, M. J. in *Asteroids II* (eds Binzel, R. P. et al.) 921–948 (Univ. Arizona Press, Tucson, 1989).
20. Gaffey, M. J., Burbine, T. H. & Binzel, R. P. *Meteoritics* **28**, 161–187 (1993).
21. Gaffey, M. J. et al. *Icarus* **106**, 573–602 (1993).
22. Lipschutz, M. E., Gaffey, M. J. & Pellas, P. in *Asteroids II* (eds Binzel, R. P. et al.) 740–777 (Univ. Arizona Press, Tucson, 1989).
23. Wasson, J. T. *Meteorites* 175–176 (Springer, New York, 1974).
24. Carrier, W. D., Olhoeft, G. R. & Mendell, W. in *Lunar Source Book* (eds Heiken, G. et al.) 475–594 (Cambridge Univ. Press, 1991).

ACKNOWLEDGEMENTS. Measurements of the positions of Ida were contributed by B. E. A. Mueller and H. Breneman. We thank the Director of the Space Telescope Science Institute for assistance in acquiring observing time on the Hubble Space Telescope at short notice, and collaborators E. N. Wells and D. Tholen. We acknowledge the perspicacity of A. Harch who first noticed the presence of Dactyl in the SSI data. We acknowledge the support and encouragement of the entire Galileo Imaging Team in this work. The National Optical Astronomy Observatories are operated by AURA, Inc. for the US NSF. This work was supported by NASA through the Jet Propulsion Laboratory Galileo Project.

Magnetic switching in cobalt films by adsorption of copper

W. Weber*, C. H. Back†, A. Bischof*, D. Peschl† & R. Allenspach*

* IBM Research Division, Zurich Research Laboratory, CH-8803 Rüschlikon, Switzerland

† Laboratorium für Festkörperphysik, ETH Hönggerberg, CH-8093 Zürich, Switzerland

MAGNETIC storage of information requires the ability to manipulate the magnetization of thin films with high sensitivity and spatial resolution. Non-magnetic overlayers are known to affect the characteristics of magnetic films: for example, the direction of magnetization of cobalt and iron films can be altered by deposition of a monolayer of copper and gold, respectively^{1–3}. The magnetic properties of cobalt films seem to be particularly sensitive to copper overlayers—deposition of only sub-monolayer amounts of copper will decrease the coercive field required to invert the magnetization direction⁴. Here we show that copper coverages as small as three-hundredths of a monolayer are sufficient to rotate by 90° the magnetization of Co films up to 20 atomic layers thick. This implies that the spins of about 500 cobalt atoms switch direction for each copper atom added. Adding more copper eventually switches the magnetization back to its original direction. This fine tuning of thin-film magnetism might be useful for developing sensitive magnetic-field sensors, as well as for magnetic recording.

We use Co films grown epitaxially on a Cu(100) single crystal with a slight miscut of 1.6° in the [110] direction—this produces a controlled step structure. These steps turn out to play a central part in the magnetic switching. To detect changes in the magnetization, we use spin-polarized scanning electron microscopy (spin-SEM)⁵ and measurements of the magneto-optical Kerr effect⁶. Spin-SEM determines the magnetization direction with high precision and spatial resolution, and the Kerr effect records hysteresis loops in the magnetization as a function of applied magnetic field.

To investigate the changes that might take place on Cu deposition, we grew a Cu wedge of 120 μm length onto a Co film and measured the magnetization direction locally by spin-SEM as a

function of the Cu coverage d_{Cu} . Figure 1 shows a magnetic map of a 2.16-nm-thick Co film (≈ 12 atomic layers, AL) partly covered by such a Cu wedge. The entire image, except one stripe, appears dark. This is direct evidence that the magnetization vector of the uncovered as well as the Cu-covered parts ($d_{\text{Cu}} > 1.2$ AL) of the Co film point in the [110] direction, which runs parallel to the steps induced by the miscut of the substrate⁷. In contrast, at low Cu coverages the magnetization direction changes in-plane by 90° from the [110] to the [110] direction, (henceforth denoted $\parallel$ and $\perp$, respectively), as indicated by the light stripe. A line scan through the magnetic map in Fig. 1 reveals that this transition of the magnetization direction by 90° is discontinuous. This abruptness, and the fact that the original magnetization direction is resumed on further Cu coverage, signifies the elementary features of a magnetic switch: the magnetization switches discontinuously between two discrete states.

To elucidate the mechanism for this behaviour, we investigated the response of the magnetization to a magnetic field. Hysteresis loops were recorded using the Kerr effect on increasing the Cu coverage of a Co film (Fig. 2). Again, the easy magnetization axes of the uncovered and the Cu-covered ($d_{\text{Cu}} > 1.2$ AL) Co films point in the $\parallel$ -direction. This is clear from the very different hysteresis loops observed for the two directions: a rectangular loop typical of the easy magnetization axis for the $\parallel$ -direction, and a more complicated one composed of two shifted loops for the $\perp$ -direction. Correspondingly, the ratio of the remanent magnetization $M_r \equiv M(H=0)$ to the saturation magnetization $M_s \equiv M(H=14 \text{ kA m}^{-1})$ is $M_r/M_s = 1$ for the $\parallel$ -direction but $M_r/M_s = 0$ for the $\perp$ -direction.

The region of interest has a Cu coverage of 0–1.2 AL. Minute amounts of Cu coverage completely change the shape of hysteresis loops. For coverages thicker than 0.03 AL, the role of the two directions is interchanged. The complicated loop of the $\perp$ -direction is replaced by the rectangular loop of an easy magnetization axis, whereas the $\parallel$ -direction loses its easy character, and $M_r/M_s < 0.2$. An additional 0.1 AL of Cu switches the easy magnetization axis back to the $\parallel$ -direction, and produces rectangular hysteresis. Most intriguing is the $\perp$ -loop at $0.15 \text{ AL} < d_{\text{Cu}} < 1.2 \text{ AL}$. Although $M_r/M_s \approx 1$, the typical rectangular loop of an easy axis is missing. A compound loop is present, but the shift of the loops is smaller than that of the uncovered Co film in $\perp$ -direction. This means that because $M_r/M_s \approx 1$ for both

directions, the magnetization is pinned in the direction in which the field was applied before. This metastable behaviour of the $\perp$ -direction disappears on further Cu coverage as soon as the loop shift increases to $M_r/M_s < 1$.

The observed switching phenomenon must be attributed to the preferential step direction. The breaking of the four-fold symmetry at a step edge leads to the uniaxial magnetic anisotropy shown by the hysteresis loops on the bare Co film. Scanning tunnelling microscopy (STM) verifies that the step arrangement of the Cu substrate is largely replicated by the epitaxial Co film. (A. Vaterlaus *et al.*, unpublished observations). Moreover, STM reveals that the adsorbed Cu atoms decorate the steps. We expect this decoration to cause the magnetization switching by a strong change of the magnetic anisotropy at the step sites, induced by hybridization of the Cu and Co wavefunctions. This model implies that the switching coverage normalized to a completed layer, d_s , should be proportional to the density of steps, that is $d_s = \tan \alpha$, where α is the miscut angle. In fact, $d_s = 0.03 \approx \tan \alpha$ for $\alpha = 1.6^\circ$. Additional experiments on a Cu substrate with a miscut of $\alpha = 3.4^\circ$ yield $d_s = 0.07$, in agreement with the simple model. It is clear that this model only holds for a limited range of angles. If α is too small, the average step distance is too large for the Co spins on the terrace to follow the switch of the spin at a Co step site. On the other hand, the second switch of the easy magnetization axis occurs at larger Cu thickness, where the terrace sites become dominant. In fact the magnetic anisotropy saturates at ≈ 1 AL, relating this switching to hybridization^{1,2} at terrace sites.

We stress that the magnetization direction of the Co films switches throughout the entire thickness. Hence the magnetic properties of the entire film are governed by the changes induced at the surface, specifically at the steps. Viewed on an atomic scale, the spin at the Co site rotates 90° on Cu adsorption. The exchange energy then aligns the neighbouring Co spins with the step-site spin and thereby transports the influence of the step into the terrace, provided the energy barrier against such a change is overcome. Considering the magnetic energy of the ensemble of the Co film and adsorbate, this means that per single Cu atom, on average, the magnetization direction of more than 500 Co atoms must switch.

This effect offers new prospects for studying magnetism in general. For instance, an adsorbed Cu atom could possibly be

used as a very sensitive experimental probe of the energetics at steps. At present, it is a formidable task to calculate magnetic anisotropies *ab initio* on a perfect surface⁸, and virtually impossible to estimate reliably the influence of "imperfections" such as steps. Measurements of the kind presented here can provide experimental input for such calculations. Such measurements might also be used in experiments to determine one of the most fundamental and yet unexplored quantities in magnetism, the correlation length. By adsorbing Cu at a step site, the Co spin at this site is forced into the perpendicular direction, and the decay length of this disturbance is determined via the minimum miscut angle α for which all spins on the terrace become reoriented.

A further class of experiments can now exploit the epitaxial engineering aspects of this technique: microstructures of controlled lateral extent, and having hitherto unexplored magnetic characteristics, can be produced. Among the possibilities are magnetically highly frustrated systems reminiscent of spin glasses, and artificial magnetic domain structures. Either of these

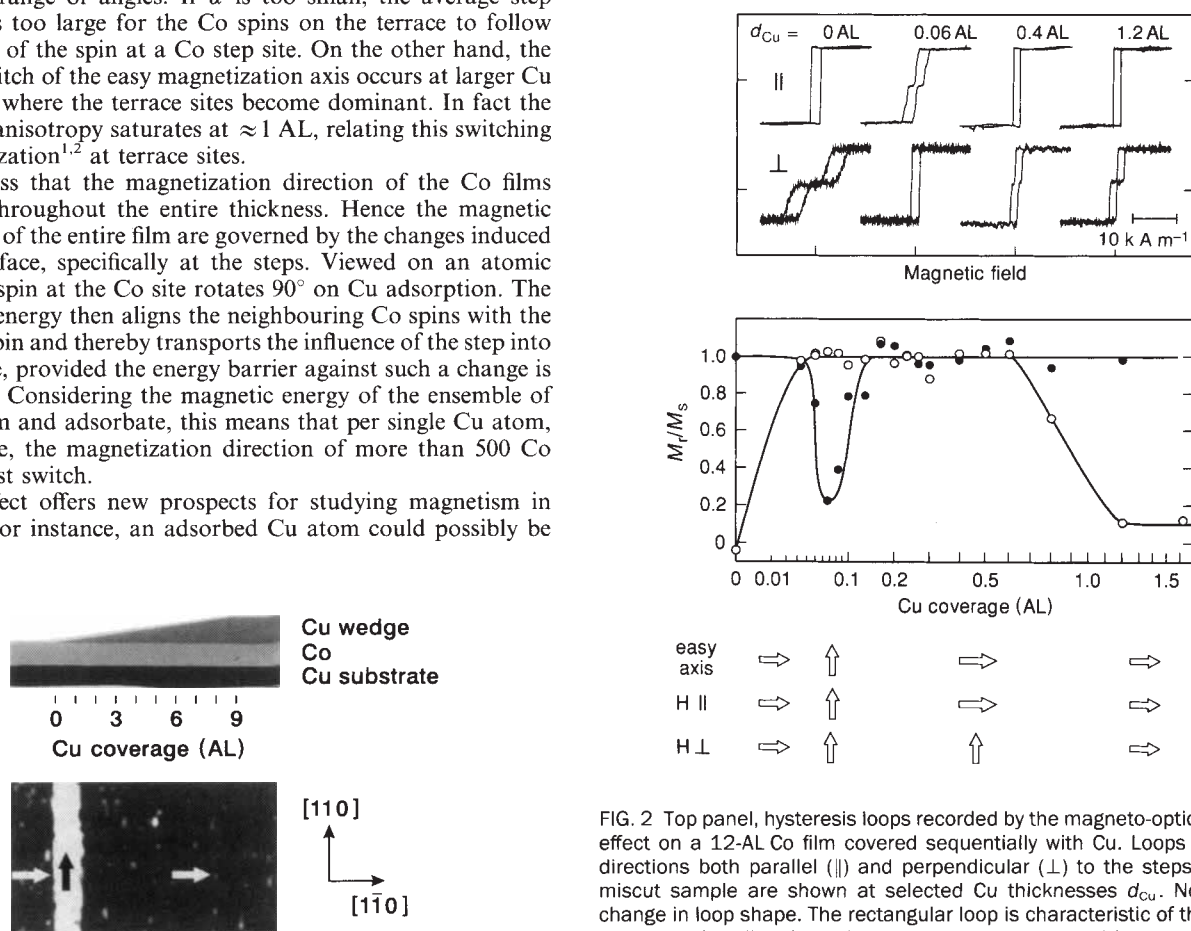


FIG. 2 Top panel, hysteresis loops recorded by the magneto-optical Kerr effect on a 12-AL Co film covered sequentially with Cu. Loops for the directions both parallel ($\parallel$) and perpendicular ($\perp$) to the steps of the miscut sample are shown at selected Cu thicknesses d_{Cu} . Note the change in loop shape. The rectangular loop is characteristic of the easy magnetization direction, whereas a more complicated loop composed of two smaller loops is observed in the other direction. Middle panel, the ratio of remanent magnetization M_r to saturation magnetization M_s as a function of Cu coverage for both directions ($\parallel$, solid circles; $\perp$, open circles; lines are to guide the eye). The d_{Cu} scale is nonlinear ($\sqrt{d_{Cu}}$) to highlight the low-coverage region of primary interest. The in-plane switching of the magnetization from $\parallel$ to $\perp$ occurs at 0.03 AL, the switching back from $\perp$ to $\parallel$ at 1.2 AL. The intermediate-thickness range between 0.15 AL and 1.2 AL is a metastable state: the remanent magnetization stays in the direction in which the field was previously applied. As the wedge sample was measured as grown in zero magnetic field, no switching of the magnetization direction at 0.15 AL was observed in the spin-SEM experiment of Fig. 1. Bottom panel, the easy magnetization axis is indicated by arrows, as is the magnetization direction after switching off the field in either the $\parallel$ - or the $\perp$ -direction.

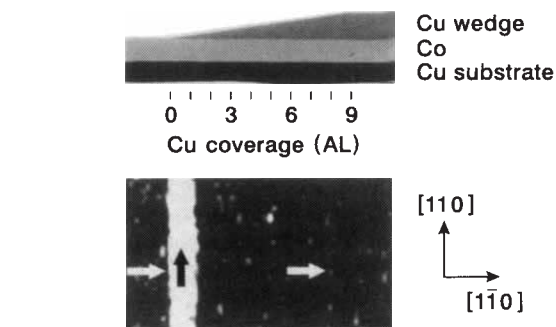


FIG. 1 Bottom, spin-SEM image ($180 \times 60 \mu\text{m}$) of a 12-AL-thick Co film on a slightly miscut Cu(100) single crystal covered with a Cu wedge (0–9 AL). Top, cross-section through the sample. The arrows on the spin-SEM image indicate the magnetization direction of the Co film; the coordinate system on the right specifies the crystallographic directions of this film. The spin-SEM image shows a sharp transition of the magnetization direction, from being parallel to the steps of the miscut specimen to being perpendicular to these steps. This discontinuity contrasts with the continuous rotation observed in other systems, such as Co/Au(111) on adsorption of Cu (ref. 1). The discrepancy has its origin in the anisotropy contributions that differ for the uncovered Co/Cu(100) and Co/Au(111) films, leading to a discontinuous and continuous magnetization rotation, respectively¹².

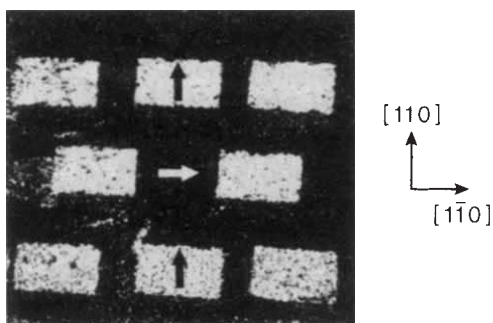


FIG. 3 Spin-SEM image of a 12-Å Co film covered by 0.06 Å of Cu through a mask having rectangular holes $200 \times 350 \mu\text{m}$ in size. The white rectangles are the covered regions, whereas the dark parts are uncovered. The arrows indicate the magnetization direction of the Co film. The coordinate system on the right indicates the crystallographic directions of the Co film. Straight 90° domain walls are induced at the boundaries between covered and uncovered regions. It is a wall type that has not been identified previously in ultra-thin magnetic films. Usually, very irregular domain walls are observed¹³. If we apply a magnetic field of appropriate direction and strength, the magnetization can be switched in a well defined part of the sample, such as the areas covered by Cu, resulting in a contrast change only in the rectangles.

could be made simply by covering the desired parts of a Co film with a few hundredths of an atomic layer of Cu. An example of this application is shown in Fig. 3. Through a mask having an array of rectangular holes, 0.06 Å of Cu has been evaporated onto the Co film. This provokes a local patterned switching of the magnetization direction only in the covered parts of the specimen.

These results might also have practical applications. For example, one possibility is the fabrication of materials having finely tuned magnetic properties by positioning Cu-covered patches of the appropriate size and shape in close vicinity to produce a magnetic device on the scale of micrometres. Magnetic-field sensors, such as the ones employed in magnetic storage devices, could profit from the fact that a means now exists to reorient the magnetization intrinsically by 90° without the need for additional external biasing. This is exactly the orientation that is required for the efficient functioning of a giant magnetoresistive head⁹, which is expected to become the state-of-the-art readout element in the foreseeable future.

The ultimate goal would certainly be to investigate the influence of a single Cu atom when brought into close proximity to the spins of neighbouring Co atoms. Such experiments have recently been performed using STM techniques¹⁰ for Fe atoms on a Cu surface¹¹. A magnetic version of these experiments, however, has yet to be performed. □

Molecular redox switches based on chemical triggering of iron translocation in triple-stranded helical complexes

Lior Zelikovich, Jacqueline Libman & Abraham Shanzer*

Department of Organic Chemistry, The Weizmann Institute of Science, Rehovot 76100, Israel

* To whom correspondence should be addressed

THE growing interest in miniaturization of electronic components is stimulating research on molecular assemblies with device-like functionalities^{1–5}. Molecule-based devices have been reported that might act as sensors^{6–8}, diodes^{9–11}, logic gates¹² and switches^{5,13–19}. Molecular switches should ideally be able to respond controllably and reversibly to external triggers^{7,12,14,15,20,21}. Here we report the synthesis of molecular redox switches based on helical metal complexes^{22–29} in which an iron ion can occupy one of two distinct binding cavities. Reversible translocation of the metal ion between these sites is achieved by chemical oxidation and reduction, owing to the different coordination preferences of the Fe(II) and Fe(III) states, and can be readily monitored spectroscopically.

The switches introduced here are based on triple-stranded, helical metal complexes^{22,23} that accommodate a single metal ion in either their internal, 'hard' binding cavity or their external, 'soft' cavity (Fig. 1). Chemical reduction of the guest ions from their higher to their lower oxidation states, that is from Fe(III) to Fe(II), induces the metals' translocation from the internal hydroxamate binding sites to the external bipyridyl sites with a simultaneous change of colour from light-brown ($\lambda_{\text{max}} = 420 \text{ nm}$, characteristic of Fe(III)-hydroxamate complexes)³⁰ to deep purple ($\lambda_{\text{max}} = 540 \text{ nm}$, characteristic of Fe(II)-bipyridyl complexes)³¹. Subsequent oxidation reverses this process. The distinct spectral properties of the two states allows quantitative monitoring of the switching process, while their C_3 -symmetric topology enables their incorporation into monolayers³². Moreover, the modular assembly of these switches enables optimization of their performance by systematic modifications of their symmetry, the nature and location of their ion binding sites, and

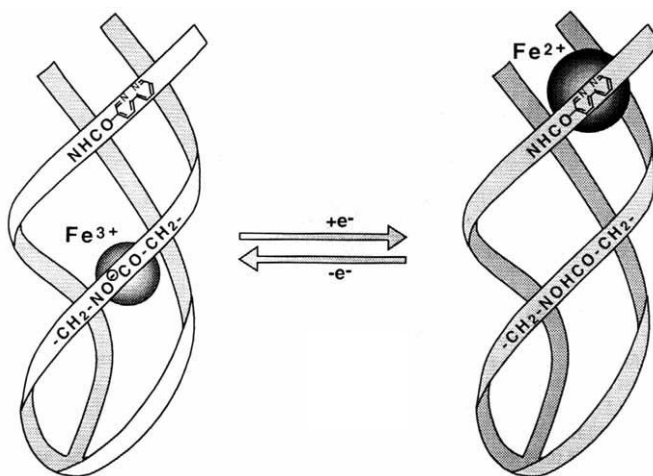


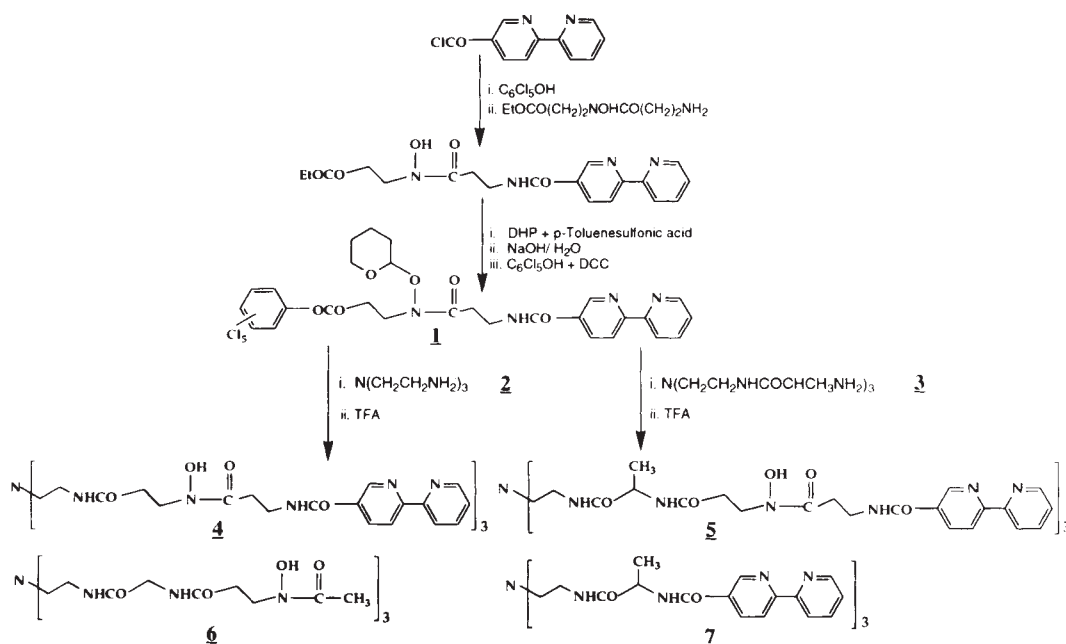
FIG. 1 Schematic representation of the triple-stranded metal complexes that function as molecular switches by interconverting the light-brown Fe(III)-complex and the purple Fe(II)-complex on exposure to reducing and oxidizing agents, respectively.

Received 30 November 1994; accepted 14 March 1995.

- Ould-Mahfoud, S. et al. *Mater. Res. Soc. Symp. Proc.* **313**, 251–256 (1993).
- Engel, B. N., Wiedmann, M. H., Van Leeuwen, R. A. & Falco, C. M., *Phys. Rev.* **B48**, 9894–9897 (1993).
- Elmers, H. J. & Gradmann, U. *Surf. Sci.* **304**, 201–207 (1994).
- Schumann, F. O., Buckley, M. E. & Bland, J. A. C., *J. appl. Phys.* **76**, 6094–6095 (1994).
- Allenspach, R. *Physics World*, **7**, 44–49 (1994).
- Bader, S. D. J. *Magn. magn. Mater.* **100**, 440–454 (1991).
- Krams, P., Hillebrands, B., Guentherodt, G. & Oepen, H. P. *Phys. Rev.* **B49**, 3633–3636 (1994).
- Li, C., Freeman, A. J., Jansen, H. J. F. & Fu, C. L. *Phys. Rev.* **B42**, 5433–5442 (1990).
- Tsang, C. et al. *IEEE Trans. Mag.* **30**, 3801–3806 (1994).
- Eigler, D. M. & Schweizer, E. K. *Nature* **344**, 524–526 (1990).
- Crommie, M. F., Lutz, C. P. & Eigler, D. M. *Science* **262**, 218–220 (1993).
- Grolier, V., Ferré, J., Maziewski, A., Stefanowicz, E. & Renard, D. *J. appl. Phys.* **73**, 5939–5941 (1993).
- Oepen, H. P., Benning, M., Ibach, H., Schneider, C. M. & Kirschner, J. J. *Magn. magn. Mater.* **86**, L137–L142 (1990).

ACKNOWLEDGEMENTS. We thank F. O. Schumann and J. A. C. Bland for communicating their results before publication. This work was supported by KWF and BBW within the HCM programme.

FIG. 2 Synthetic scheme for the preparation of the ditopic triple-stranded ligands **4** and **5**, and structures of the monotopic reference ligands, trishydroxamate **6** and trisbipyridine **7**.



the choice of the redox-active metal ions. A change in coordination preference of metal ions induced by oxidation and reduction allowed Livoreil *et al.*²¹ to develop an electrochemically switchable catenane.

The synthesis of our redox switches (Fig. 2) involved preparation of the single-stranded chain **1** possessing both ion binding groups (bipyridyl and hydroxamate), and condensation of this chain **1** with the tripodal anchors, either the achiral tris(aminoethyl) amine **2** or the chiral trisalanil derivative **3**, similarly to the procedures earlier employed²³. Removal of the protecting groups by treatment with trifluoroacetic acid provided the desired products **4** and **5**, whose spectroscopic properties such as Fourier transform infrared, ¹H- and ¹³C-NMR and electrospray mass spectroscopy (A.-M. Albrecht-Gary, A. van Dorselaer, A.S., E. Leize and L.Z., unpublished results) were in agreement with the assigned structures. Titration (monitored by ultraviolet-visible and circular dichroism (CD) spectroscopies) of the chiral derivative **5** with Fe(III) led to the left-handed Fe(III)-hydroxamate complex (Table 1)^{23,33}.

Treatment of the chiral complex **5**-Fe(III) with ascorbic acid³⁴ resulted in rapid reduction, as evidenced by the appearance of the purple-red absorption of the **5**-Fe(II)-bipyridyl complex

(Table 1 and Fig. 3). The spectral features of the latter were identical with those of the **5**-Fe(II)-bipyridyl complex obtained directly by treatment of the ligand **5** with FeSO₄. Reaction of **5**-Fe(II) for a few minutes with ammonium persulphate at 70 °C (ref. 35) caused reversal of this process, as manifested by the disappearance of the purple-red colour and regeneration of the light-brown colour of **5**-Fe(III) (Fig. 3). These redox processes are most likely to involve intra-molecular ion translocation, as the analogous inter-molecular processes between the monotopic Fe(III)-trishydroxamate complex **6**-Fe(III) and bipyridine or trisbipyridine **7** failed to occur under identical conditions. The achiral redox-switch **4**-Fe(III), which differs from the chiral redox-switch **5**-Fe(III) by lacking its L-alanyl acid residue, followed qualitatively the same behaviour, but differed in the rate of the reduction. The half-life time *t*_{1/2} of the reduction process was ~45 s for the chiral switch **5**-Fe(III), and was 15 s for the achiral switch **4**-Fe(III). The different behaviour of these two switches illustrates the pronounced effect of structural variations, and indicates the advantage of modular design and synthesis in facilitating modifications in structure and performance.

The orientation of the bipyridyl chromophores is significant. These chromophores are randomly oriented in the free ligand **5**,

TABLE 1 Spectra of ligands and their iron complexes

Compound	UV $\lambda_{\text{max}}(\epsilon \cdot 10^{-4})$	CD $\lambda_{\text{ext}}(\Delta \epsilon)$	UV $\lambda_{\text{max}}(\epsilon \cdot 10^{-4})$	CD $\lambda_{\text{ext}}(\Delta \epsilon)$
5 free ligand*†	292 (6.33)			
5 -Fe(III) complex*‡	292 (6.35)	281 (+3.93), 291 (0), 302 (-7.24)	420 (0.24)	370 (-0.93), 417 (0), 453 (+0.53)
5 -Fe(II) complex*§	301 (6.35)	286 (+11.8), 293 (0), 306 (-25.2)	542 (0.47)	470 (-1.26), 505 (0), 545 (+2.25)
5 -Fe(III)+Fe(II) complex	301 (6.40)	285 (+18.8), 296 (0), 307 (-44.8)	542 (0.55)	475 (-1.79), 505 (0), 545 (+3.80)
7 -Fe(II) complex	298 (6.64)	287 (+10.1), 295 (0), 308 (-18.5)	551 (0.51)	470 (-1.13), 520 (0), 547 (+1.62)
6 -Fe(III) complex			423 (0.23)	

All spectra were measured at concentrations of 0.2–0.3 mM in methanol: aqueous 0.1 N Tris-buffer (9:1), final pH 7.5. Symbols and units: λ_{max} , wavelength of maximal absorption (nm); ϵ , molar absorption coefficient ($10^3 \text{ cm}^2 \text{ M}^{-1}$); λ_{ext} , wavelength of extreme circular dichroism (nm); $\Delta \epsilon = \epsilon_L - \epsilon_R$, difference of absorption coefficient between left and right circularly polarized light (10^3 cm^2).

* A.-M. Albrecht-Gary, A. van Dorselaer, A.S., E. Leize and L.Z., unpublished results.

† Electrospray mass spectrum: $[\mathbf{5} + \text{H}]^+$ ($m/z = 1380.1$), $[\mathbf{5} + 2\text{H}]^{2+}$ ($m/z = 690.5$); require 1,380.3 and 690.1, respectively.

‡ Electrospray mass spectrum: $[\mathbf{5} + \text{Fe} + 3\text{H} + \text{H}]^+$ ($m/z = 1433.7$), $[\mathbf{5} + \text{Fe} + 3\text{H} + 2\text{H}]^{2+}$ ($m/z = 717.1$); require 1,433.3 and 717.0, respectively.

§ Electrospray mass spectrum: $[\mathbf{5} + \text{Fe} + \text{Cl}]^+$ ($m/z = 1471.3$), $[\mathbf{5} + \text{Fe}]^{2+}$ ($m/z = 718.5$); require 1,471.3 and 718.1, respectively.

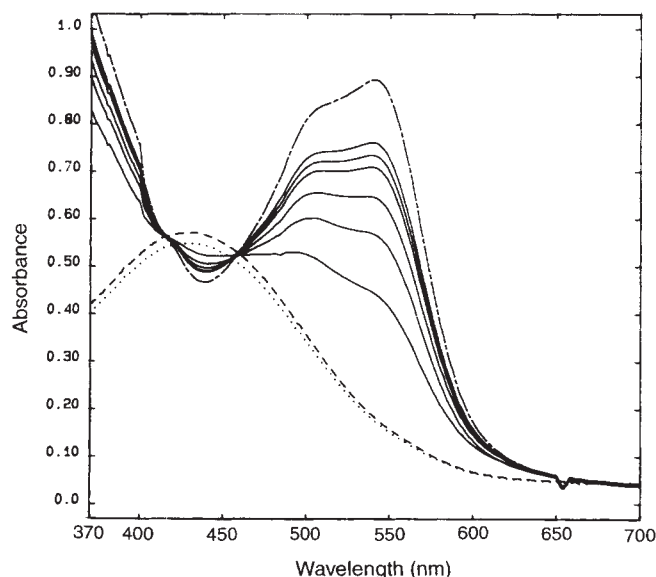


FIG. 3 Ultraviolet-visible absorption spectra of **5**-Fe(III) (dashed line) and **5**-Fe(II) (dot-dashed line), (0.25 mM concentration in MeOH:aqueous Tris buffer (9:1 ratio), pH 7.5); also shown are changes of the spectrum of **5**-Fe(III) (solid lines) on treatment with ascorbic acid (fivefold excess, 15-s intervals), and the spectrum (dotted line) after treatment with ammonium persulphate (tenfold excess).

but become increasingly aligned in a chiral sense when the internal **5**-Fe(III) complex is converted to the external **5**-Fe(II) complex and the dinuclear **5**-Fe(III)/Fe(II) complex. (This dinuclear complex was prepared by treatment of **5** with one equivalent of FeSO₄ and one equivalent of FeCl₃. The absorption spectrum of the dinuclear complex in the visible region is a superposition of those obtained from the corresponding mononuclear complexes **5**-Fe(II) and **5**-Fe(III), and lacks new charge-transfer bands. This excludes the occurrence of metal-metal interactions.) This is shown by the circular dichroism spectra of these complexes which reveal increasing Cotton effects in the 290-nm region due to exciton coupling between the bipyridyl groups (Table 1)³⁶. It may accordingly be concluded that the internal metal centre directs and amplifies the chirality of the external one. Moreover, the reductive switching process seems to occur diastereoselectively, as the alignment of the bipyridyl chromophores intensifies during the reduction of **5**-Fe(III) to **5**-Fe(II). On re-oxidation, the internal **5**-Fe(III) complex is regenerated, although equilibration to its original left-handed configuration is only established after several hours. The rather slow regeneration of the left-handed **5**-Fe(III) complex is probably due to the rather drastic oxidation conditions; but the possibility that the oxidized Fe(III) is solvated before its recapture by the hydroxamate ligand cannot be excluded.

The molecular switches introduced here are intended to contribute to the future development of molecule-based technologies, although significant improvements in their performance are still needed. The challenges that remain to be addressed include developing a means to stimulate switching by electrochemical and photochemical means, improving their addressing, reading and response rate, and finding ways of immobilizing them on conducting surfaces. □

Received 16 December 1994; accepted 22 March 1995.

1. Carter, F. L. *Molecular Electronic Devices II* (Dekker, New York, 1987).
2. Hopfield, J. J., Onuchic, J. N. & Beratan, D. N. *J. phys. Chem.* **93**, 6350–6357 (1989).
3. Lehn, J. M. *Angew. Chem. int. Edn engl.* **29**, 1304–1319 (1990).
4. Wild, U. P., Bernert, S., Kohler, B. & Renn, A. *Pure appl. Chem.* **64**, 1335–1342 (1992).
5. Anders, J. et al. *Ber. Bunsenges. phys. Chem.* **97**, 483–487 (1993).
6. Rubinstein, I., Steinberg, S., Tor, Y., Shanzer, A. & Sagiv, J. *Nature* **332**, 426–429 (1988).

7. Huston, M. E., Akkaya, E. U. & Czarnik, A. W. *J. Am. chem. Soc.* **111**, 8735–8737 (1989).
8. Bissell, R. A. et al. *Chem. Soc. Rev.* **21**, 187–195 (1992).
9. Aviram, A. & Ratner, M. A. *Chem. Phys. Lett.* **29**, 277–283 (1974).
10. Pomerantz, M., Aviram, A., McCorkle, A., Li, L. & Schrott, A. G. *Science* **255**, 1115–1118 (1992).
11. Martin, A. S., Sambles, J. R. & Ashwell, G. J. *Phys. Rev. Lett.* **70**, 218–221 (1993).
12. de Silva, A. P., Gunaratne, H. Q. N. & McCoy, C. P. *Nature* **364**, 42–44 (1993).
13. Feringa, B. L., Jager, W. F. & de Lange, B. J. *Am. chem. Soc.* **113**, 5468–5470 (1991).
14. Wasielewski, M. R., O'Neill, M. P., Gosztola, D., Niemczyk, M. P. & Svec, W. A. *Pure appl. Chem.* **64**, 1319–1325 (1992).
15. Gilat, S. L., Kawai, S. H. & Lehn, J. M. *J. chem. Soc., chem. Commun.* 1439–1442 (1993).
16. Goulle, V., Harriman, A. & Lehn, J. M. *J. chem. Soc., chem. Commun.* 1034–1036 (1993).
17. Voegtli, F., Mueller, W. M., Mueller, U., Bauer, M. & Rissanen, K. *Angew. Chem. int. Edn engl.* **32**, 1295–1297 (1993).
18. Bissell, R. A., Cordova, E., Kaifer, A. G. & Stoddart, J. F. *Nature* **369**, 133–137 (1994).
19. Joulie, L. F., Schatz, E., Ward, M. D., Weber, F. & Yellowlees, L. J. *J. chem. Soc., Dalton Trans.* 799–804 (1994).
20. Aviram, A. *Int. J. Quantum Chem.* **42**, 1615–1624 (1992).
21. Livoreil, A., Dietrich-Buchecker, C. O. & Sauvage, J.-P. *J. Am. chem. Soc.* **116**, 9399–9400 (1994).
22. Tor, Y. *Artificial Tripodal Ligands: Design, Synthesis and Properties* (Weizmann Inst. of Science, Rehovot, Israel, 1990).
23. Libman, J., Tor, Y. & Shanzer, A. *J. Am. chem. Soc.* **109**, 5880–5881 (1987).
24. Kraemer, R., Lehn, J.-M., Cian, A. D. & Fischer, J. *Angew. Chem. int. Edn engl.* **32**, 703–705 (1993).
25. Lehn, J.-M. et al. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2565–2569 (1987).
26. Williams, A. F., Piguet, C. & Bernadinelli, G. *Angew. Chem. int. Edn engl.* **30**, 1490–1492 (1991).
27. Piguet, C., Hopfgartner, G., Bocquet, B., Schaad, O. & Williams, A. F. *J. Am. chem. Soc.* **116**, 9092–9102 (1994).
28. Constable, E. C. *Angew. Chem. int. Edn engl.* **30**, 1450–1451 (1991).
29. Constable, E. C., Hannon, M. J. & Tocher, D. A. *Angew. Chem. int. Edn engl.* **31**, 230–232 (1992).
30. Raymond, K. N., Mueller, G. & Matzkanke, B. F. *Top. Curr. Chem.* **123**, 49–102 (1984).
31. Hawker, P. N. & Twigg, M. V. (eds) Wilkinson, G., Gillard, R. D. & McCleverty, J. A. 1179–1288 (Pergamon, Oxford, 1987).
32. Gafni, Y., Weizman, H., Libman, J., Shanzer, A. & Rubinstein, I. *J. Am. chem. Soc.* (submitted).
33. van der Helm, D., Baker, J. R., Eng-Wilmot, D. L., Hossain, M. B. & Loghry, R. A. *J. Am. chem. Soc.* **102**, 4224–4231 (1980).
34. Emery, T. & Neillands, J. B. *J. Am. chem. Soc.* **82**, 3659–3662 (1960).
35. Burgess, J. & Prince, R. H. *J. chem. Soc., A* 1772–1775 (1966).
36. Saito, Y. in *Topics in Stereochemistry* Vol. 10 (eds Eliel, F. L. & Allinger, N. L.) 95–174 (Wiley, New York, 1978).

ACKNOWLEDGEMENTS. This work was supported by the Israel Science Foundation and the Consortium of German Chemical Companies. A.S. is the Siegfried and Irma Ullmann Professor at the Weizman Institute of Science.

Spontaneous assembly of a hinged coordination network

Geoffrey B. Gardner*, D. Venkataraman†, Jeffrey S. Moore†‡ & Stephen Lee*‡

* Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109, USA

† Departments of Chemistry, and Materials Science and Engineering, University of Illinois, Urbana, Illinois 61801, USA

THE field of supramolecular chemistry has advanced to a stage at which it is possible to select building blocks that will self-assemble into structures with specific network topologies^{1–3}. This makes possible the rational design and synthesis of molecular solids with potentially interesting properties. Here we report the construction of open, hinged networks from molecular building blocks. This class of materials has been predicted to exhibit unusual mechanical properties, including auxetic behaviour (negative Poisson's ratio) and negative coefficients of thermal expansion^{4–6}. Our approach relies on the notion that rigid organic molecules of high symmetry will adopt one of only a few possible structures when linked via hydrogen bonds or coordination to metals^{7–9}. We use trigonal ligands to make networks joined at the vertices by metal ions; the resulting networks are homeotypic¹⁰ with the honeycomb-like AlB₂ and the hinge-like ThSi₂ phases. The hinge-like network has channels of inner diameter 15 Å, within which included molecules can be exchanged while the framework remains intact. We have not yet determined whether this material is auxetic.

‡ To whom correspondence should be addressed.

Structures based on three-connected nets such as AlB_2 and ThSi_2 have evoked wide attention^{4,11–13}. Particular networks in this family are predicted to have unusual mechanical, thermal and electrical properties, especially when they contain large open features. Although syntheses of fragments of these prototypes have been achieved, extensions to three-dimensional networks by conventional synthesis have been difficult to realize^{11,14}. One approach to the rational design of crystalline solids has been to use organic molecules which, by their coordinating propensities, their geometry and their relative stoichiometries, are naturally predisposed to form one of several well categorized solid structure types. Structures recently prepared by this method include molecular analogues of the diamond, rutile, simple cubic and the PtS structures^{9,15,16}.

Following this building-block approach^{17–19}, we can imagine constructing the AlB_2 or ThSi_2 structures by replacing the parent atoms in these lattices with ligands and metal ions whose bonding propensities favour local trigonal geometry. The exact analogues for such a replacement are the LaPtSi (a ThSi_2 derivative)²⁰ and the CaCuP (an AlB_2 derivative)²¹ structure types. The 'noble' metal (Pt or Cu) atoms would be replaced by a three-coordinate metal atom, and the main-group (Si or P) atoms by the three-fold symmetric tritopic ligand. The electro-positive La or Ca atom positions would be replaced by a non-coordinating anionic counter-ion. Trigonal geometries for the tritopic ligands could be realized in the 1,3,5-functionalized benzene units. We chose $\text{Ag}(\text{I})$ as the metal ion, as it is known to exhibit three-fold coordination. The counter-ions were chosen to be the weakly coordinating triflate (trifluoromethanesulphonate, CF_3SO_3^-).

The simplest tritopic molecule to consider is 1,3,5-triazine. However, molecular models based on the LaPtSi and CaCuP lattices using triazine building blocks and typical $\text{Ag} \cdots \text{N}$ distances of 2.2 Å indicate that these hypothetical networks are too crowded. Hence higher homologues were chosen as ligands. The trigonal ligands 1,3,5-tricyanobenzene (TCB) (**1**) and 1,3,5-tris(4-ethynylbenzonitrile)benzene (TEB) (**2**) were synthesized according to well established procedures^{22,23}. Nitriles were chosen as the ligating groups as they readily coordinate to the soft $\text{Ag}(\text{I})$ cation in a low-coordinate fashion. Crystals of $[\text{Ag}(\text{TCB})(\text{CF}_3\text{SO}_3)]$ and $[\text{Ag}(\text{TEB})\text{CF}_3\text{SO}_3]$ suitable for X-ray analysis were obtained by mixing the constituent molecules in benzene and cooling from elevated temperatures.

TABLE 1 Crystal data for compound **3** and **4** at 27 °C

	TCB-AgOTf (3)	TEB-AgOTf (4)
Crystal system	Trigonal	Orthorhombic
Space group	$P3_2$ (145)	$Pnn2$ (34)
Cell parameters	$a=b=10.582$ (2) Å $c=10.665$ (3) Å $\alpha=\beta=90^\circ$ $\gamma=120^\circ$	$a=11.592$ (1) Å $b=19.099$ (2) Å $c=38.749$ (4) Å $\alpha=\beta=\gamma=90^\circ$
Z	3	4
Volume	1034.2 (3) Å ³	8584.1 (15) Å ³
No. of reflections refined	2,726	6,448
No. of parameters	208	313
Unweighted agreement factor (R_1)	0.036 ($F > 4\sigma$) 0.043 (all data)	0.137 ($F > 4\sigma$) (isotropic) 0.235 (all data) 0.1063 (for 2,490 $> 4\sigma$ reflections for $2\theta < 35^\circ$)
Weighted agreement factor (weighted R_2)	0.0798 (all data)	0.2867 (all data)
Highest peak in final difference map	0.0770 ($F_0 > 4\sigma$) 0.26 e Å ⁻³	0.2644 ($F_0 > 4\sigma$)* 0.54 e Å ⁻³
Goodness of fit	1.04	1.64

Experimental details for **3**: diffractometer, Syntex $P2_1$ IV; $\lambda(\text{Mo K}\alpha)=0.710734$ Å, Lp corrected; 8,153 reflections collected; computing structure solution, SHELXTL PLUS; computing structure refinement, SHELXL-93²⁹; reflections were refined based on F_0^2 by full matrix least squares. Experimental details for **4**: diffractometer, Siemens SMART system with a CCD (charge-coupled device) detector; $\lambda(\text{Mo K}\alpha)=0.710734$ Å, Lp corrected; 6,448 reflections collected; structure solution and refinement, as for **4**. For original data, see Supplementary Information.

* Suitable weights were chosen to obtain smaller estimated standard deviations (e.s.d.s) rather than the recommended weights by SHELXL-93.

The crystal structure of $[\text{Ag}(\text{TCB})(\text{CF}_3\text{SO}_3)]$ (**3**) is shown in Fig. 1A. It is homeotypic with the CaCuP structure type (Fig. 1B). The crystal structure is trigonal and consists of honeycomb sheets based on alternating 1,3,5-tricyanobenzene and $\text{Ag}(\text{I})$ units. Within each sheet, the silver ions are trigonally coordinated to three nitriles in an end-on arrangement at a distance

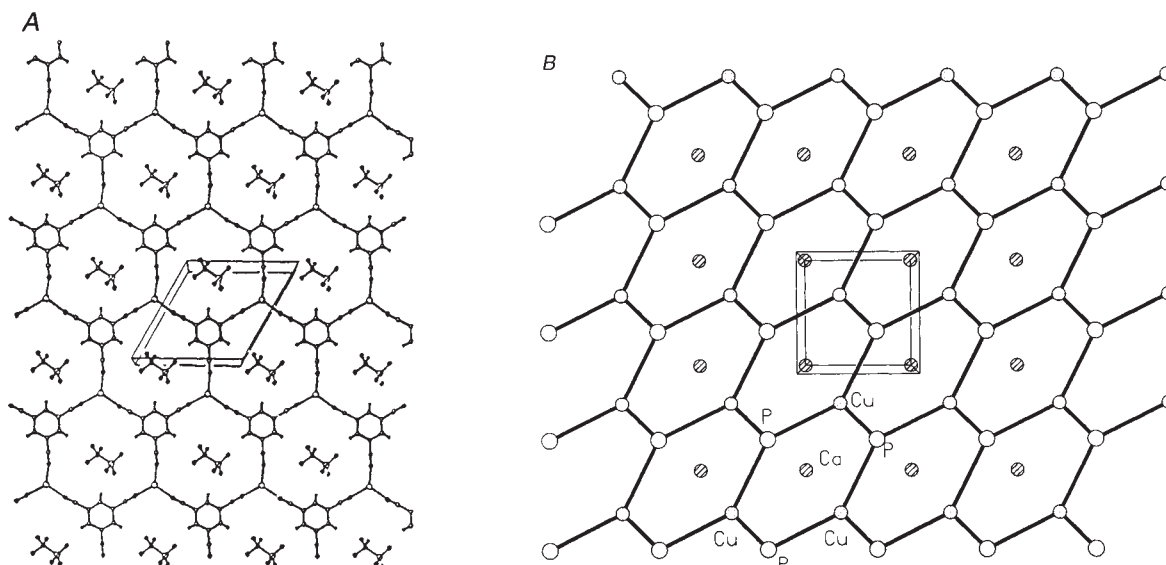


FIG. 1 A, Single-layer motif from the crystal structure of **3** parallel to the (001) plane. The triflate coordinates to the $\text{Ag}(\text{I})$ ion in the adjacent layer. B, Single-layer motif of prototype CaCuP .

of 2.261 Å (N–Ag–N angle, 114.9°; N–Ag–N angle, 122.0°). The layers are separated by 3.55 Å and repeat in an $\cdots ABCABC \cdots$ stacking sequence. The honeycomb sheets create cavities of diameter 10.03 Å. The triflate ions are weakly coordinated to the silver in the layer at the axial position of a trigonal pyramid at a distance of 2.390 Å, and consequently occupy the cavity space in an adjacent layer. The $[\text{Ag}(\text{TCB})(\text{CF}_3\text{SO}_3)]$ crystal

therefore does not have an open channel structure. The final crystal structure suggests that the method of network design by analogy to prototypic structures should be very effective.

The crystal structure of $[\text{Ag}(\text{TEB})\text{CF}_3\text{SO}_3]$ (**4**) is homeotypic with the LaPtSi structure. In contrast to the LaPtSi structure where each of the atoms are roughly comparable in size, the tritopic ligand **2** is significantly larger than either the Ag(I) or the CF_3SO_3^- ions. As a result, a single LaPtSi type net constructed with the dimensions of the tritopic ligand **2** and Ag(I) generates large void spaces. The triflate anion is of insufficient size to occupy fully these voids. The voids along [010] and [001] are filled by the six mutually independent, interpenetrating LaPtSi type lattices that constitute the $[\text{Ag}(\text{TEB})\text{CF}_3\text{SO}_3] \cdot (\text{C}_6\text{H}_6)_2$ structure (Fig. 2C)^{24–27}. We note that the interpenetration occurs in such a way as to accommodate the propensity of aromatic rings to lie in a slightly staggered coplanar arrangement with an offset angle (α) 37° at an interplanar distance (d) of ~3.3 Å (Fig. 2B). To accommodate the geometrical requirements²⁸ (α and d) for stacking of the aromatic rings, the

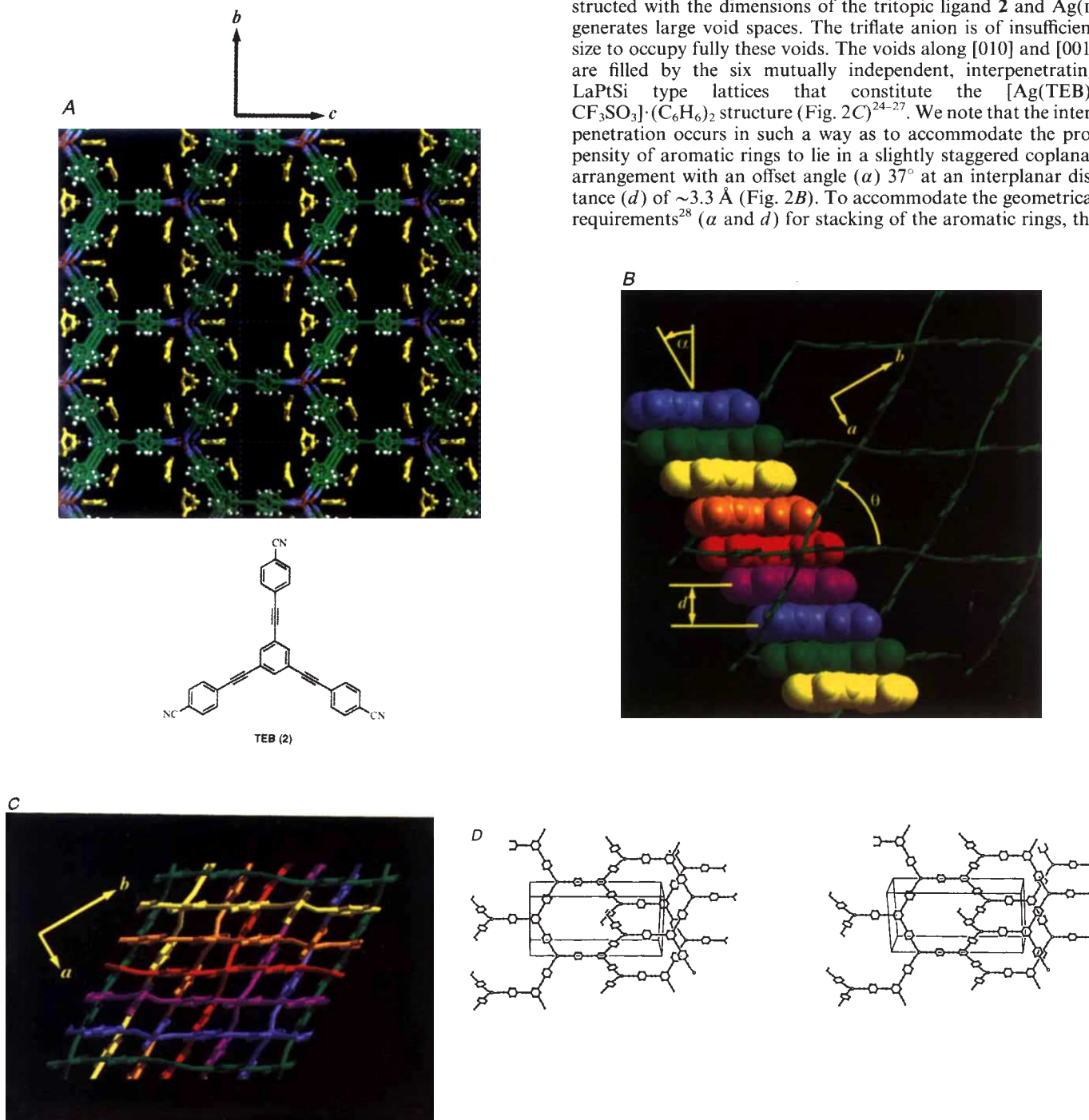


FIG. 2. A, Illustration of the 15-Å channels along [100] (*a* axis) in the crystal structure of **4**. The space within the channel is filled with benzene (shown yellow) which could readily be exchanged with C_6D_6 . The carbon atoms are shown as green, hydrogen as white, nitrogen as blue and silver as brown. The triflate counterions have been removed for clarity. The chemical structure of a molecule of TEB (**2**) is also shown. B, Illustration of the stacking of the aromatic rings in the interpenetrated

nets. Stacking occurs at a distance (d) of 3.3 Å with an offset angle α of ~37°. The shear angle (θ) of the net is 60°. This arrangement simultaneously accommodates the geometric requirements (α and d) for the six independent self-included nets. C, Six independent nets interpenetrate to fill space along [010] (*b*-axis). D, Stereoview of a single LaPtSi type net of the crystal structure of **4**, showing the molecular constituents of the framework.

nets are sheared to an angle (θ) of 60° ($\theta = 90^\circ$ for an ideal LaPtSi net, and 0° for CaCuP). However, the interpenetration leaves the channels along [100] near the maximum size possible for this network, as seen in Fig. 2A. This cavity of diameter 15 Å is filled by the solvent molecules. The coordinates of 12 benzene molecules per unit cell were located in the refinement.

Although such channels suggest a porous structure, they do not by themselves prove that the channels support reversible exchange of guest species. To demonstrate porosity, samples of $[\text{Ag}(\text{TEB})\text{CF}_3\text{SO}_3] \cdot (\text{C}_6\text{H}_6)_2$ crystals were grown from benzene. The crystals were sieved through a 120-mesh screen to remove all small crystallites. Powder X-ray diffraction studies of these crystals showed that the product was quantitatively pure, and was the same phase as the single-crystal structure. Powder-diffraction data also showed that these crystals lose their crystallinity on loss of solvent. The large crystals were then placed in C_6D_6 (isotopic purity $>99.5\%$) at room temperature, and subsequently washed with fresh C_6D_6 solution five times over the course of 3.5 days. The crystals were then placed in C_6D_6 (isotopic purity $>99.96\%$), and were washed with fresh C_6D_6 solution three times over a period of one hour. Identical samples were placed under an optical microscope, and the morphology of the crystals subjected to similar treatment were monitored as a function of time. Visual inspection showed that the sample changed very little in its overall crystalline morphology during this time period. In particular, there were no microscopically observable cracks, and the dimensions and shape of the crystal remained constant. The X-ray powder pattern of these crystals after washing with C_6D_6 (isotopic purity $>99.96\%$) showed that they remained identical to the original structure. After dissolving these crystals in fully deuterated acetone, NMR spectroscopy showed (to the limits of detection) no peak at $\delta 7.34$, thereby indicating the absence of the original benzene. Thus, although the crystal appears to be structurally unchanged, the actual composition of the crystal has changed from $[\text{Ag}(\text{TEB})\text{CF}_3\text{SO}_3] \cdot (\text{C}_6\text{H}_6)_2$ to $[\text{Ag}(\text{TEB})\text{CF}_3\text{SO}_3] \cdot (\text{C}_6\text{D}_6)_2$ without dissolution and reformation of the lattice. This demonstrates that the $[\text{Ag}(\text{TEB})\text{CF}_3\text{SO}_3] \cdot (\text{C}_6\text{H}_6)_2$ crystals are indeed porous to benzene exchange.

The examples described here provide further evidence for the efficacy of the rational design of crystalline solids by analogies with solid-state structural chemistry. We have shown that it is possible to prepare molecular-based LaPtSi and CaCuP analogues by appropriate design. These networks resemble the nets studied theoretically by Baughman and Galvão⁴. Although we cannot as yet exactly predict which of the possible structure types will form from any given set of molecular constituents, the number of potential structures can be reduced greatly. Experiments are underway to measure the auxetic properties of compound 4 and to probe the diffusion of various organic solvents into the crystals of this compound. □

20. Klepp, K. & Parthé, E. *Acta Crystallogr.* **B38**, 1105–1108 (1982).
21. Mewis, A. Z. *Naturforsch.* **33B**, 983–986 (1976).
22. Bailey, A. S., Henn, B. R. & Langdon, J. M. *Tetrahedron* **19**, 161–167 (1963).
23. Zhang, J., Pesak, D. J., Ludwick, J. L. & Moore, J. S. *J. Am. chem. Soc.* **116**, 4227–4239 (1994).
24. Wang, X., Simard, M. & Wuest, J. D. *J. Am. chem. Soc.* **116**, 12119–12120 (1994).
25. Duchamp, D. J. & Marsh, R. E. *Acta Crystallogr.* **B25**, 5–19 (1969).
26. Ermer, O. & Lindenberg, L. *Helv. chim. Acta*, **74**, 825–877 (1991).
27. Ermer, O. *J. Am. chem. Soc.* **110**, 3747–3754 (1988).
28. Gavezzotti, A. & Desiraju, G. R. *Acta Crystallogr.* **B44**, 427–434 (1988).
29. Sheldrick, G. M. *Crystal Solution Program* (Inst. für Anorg. Chemie, Göttingen, 1993).

SUPPLEMENTARY INFORMATION. Requests should be addressed to Mary Sheehan at the London editorial office of Nature.

ACKNOWLEDGEMENTS. We thank Siemens Inc. and in particular C. F. Camapana for collecting the data on 4, and J. W. Kampf for collecting the data set on $[\text{Ag}(\text{TCB})_3(\text{CF}_3\text{SO}_3)]$. J.S.M. thanks the US NSF, the NSF Young Investigator programme (1992–97), the Camille Dreyfus Teacher-Scholar Awards programme and the 3M company (non-tenured faculty awards programme) for their support. S.L. thanks the A. P. Sloan Foundation (1993–95) and the J. D. and C. T. MacArthur Foundation (1993–97) for fellowships.

Segmentation of mid-ocean ridges with an axial valley induced by small-scale mantle convection

Stéphane Rouzo, Michel Rabinowicz & Anne Briais

Groupe de Recherche de Géodésie Spatiale, CNRS UPR234 & UMR39, 14, Ave Edouard Belin, 31400 Toulouse, France

THE small-scale segmentation of mid-ocean ridges with an axial rise has been modelled by considering each ridge segment as a giant crack in the lithosphere with a tip propagating along the ridge axis^{1,2}. For ridges with an axial valley, however, this type of model fails because the lithosphere is too thick to tear³. Yet such ridges are clearly segmented, as defined by morphology, gravity and structure at the 50–100 km length scale. The ridge offsets are large^{4,5}, and vary dramatically with time. This type of segmentation is commonly related to a three-dimensional, small-scale mantle flow occurring in the partially molten asthenosphere below the ridge^{6–8}. Here we propose a model for segmentation in such ridges, in which the convective flow below the ridge favours an asymmetrical breaking of the axial-valley lithosphere. This leads to the development and separation of ridge segments in a pattern that mimics the observed geometry and temporal evolution of the segmentation of most ridges with an axial valley. If our model is correct, it implies that coupling of oceanic lithosphere to small-scale convection controls the dynamics of mid-ocean ridges.

Our model starts from the buoyant convective flow pictured in Fig. 1. In this case, convection generates rolls without any preferred orientation. At a mid-ocean ridge, however, the diverging plates induce shear at the surface. When included in the calculation, this shear forces the convective flow to orient the rolls parallel to the spreading direction (Fig. 2a). Because the flow adjusts to the boundary conditions extremely slowly, it retains the memory of the initial thermal conditions for a long time. Despite the symmetric shear, currents ascending below the ridge remain asymmetric about the spreading axis, even after 150 Myr of evolution. The continuing asymmetry in the convective pattern in turn produces an asymmetric stress field in the lithosphere. We now show how these asymmetrical stresses are able to split a straight ridge axis into segments whose lengths and offsets continue to vary for several hundred million years.

Ridges with an axial valley exhibit an average axial lithospheric thickness of 8 km (refs 9–11). The valley results from the necking of this thick lithosphere^{12,13}, stretched by large-scale horizontal stresses (20–30 MPa)¹⁴, and accreted from below by

Received 23 December 1994; accepted 3 April 1995.

1. Desiraju, G. R. *Crystal Engineering: the Design of Organic Solids* (Elsevier, New York, 1989).
2. MacDonald, J. W. & Whitesides, G. M. *Chem. Rev.* **94**, 2383–2420 (1994).
3. Lehn, J. M. *Pure appl. Chem.* **66**, 1961–1966 (1994).
4. Baughman, R. H. & Galvão, D. S. *Nature* **365**, 735–737 (1993).
5. Evans, K. E. & Hutchinson, I. J. *Nature* **353**, 124 (1991).
6. Lakes, R. *Adv. Mater.* **5**, 293–296 (1993).
7. Moore, J. S. & Lee, S. *Chem. Ind.* 556–560 (1994).
8. Zaworotko, M. J. *Chem. Soc. Rev.* **23**, 283–288 (1994).
9. Robson, R. et al. *Supramolecular Architecture* (ed. Bein, T.) Ch. 19 (ACS Symp. Ser. 499, American Chemical Soc., Washington DC, 1992).
10. Lima-de-Faria, J., Hellner, E., Liebau, F., Makovicky, E. & Parthé, E. *Acta Crystallogr.* **A46**, 1–11 (1990).
11. Diederich, F. & Rubin, R. *Angew. Chem. Int. Edn engl.* **31**, 1101–1123 (1992).
12. Hoffmann, R., Hughbanks, T., Kertész, M. & Bird, P. H. *J. Am. chem. Soc.* **105**, 4831–4832 (1983).
13. Zheng, C. & Hoffmann, R. *Inorg. Chem.* **28**, 1074–1080 (1989).
14. Wu, Z., Lee, S. & Moore, J. S. *J. Am. chem. Soc.* **114**, 8730–8732 (1992).
15. Hoskins, B. F., Robson, R. & Scarlett, N. V. *J. chem. Soc., chem. Commun.* 2025–2026 (1994).
16. Abrahams, B. F., Hoskins, B. F., Michail, D. M. & Robson, R. *Nature* **369**, 727–729 (1994).
17. Kaszynski, P., Friedli, A. C. & Michl, J. *J. Am. chem. Soc.* **114**, 601–620 (1992).
18. Simard, M., Su, D. & Wuest, J. D. *J. Am. chem. Soc.* **113**, 4696–4698 (1991).
19. Fagan, P. J., Ward, M. D. & Calabrese, J. C. *J. Am. chem. Soc.* **111**, 1698–1719 (1989).

cooling of the upper mantle. The analysis of the compensation of the axial valley topography along the Mid-Atlantic Ridge provides strong evidence that the sub-valley lithosphere behaves mainly plastically, but with a small effective elastic thickness (ref. 14): Neumann and Forsyth found that in a 12.5-km-wide zone along the rift valley, the coherence and admittance can be matched¹⁴ with an effective elastic thickness of 2 km. Such a small elastic plate thickness is valid only near the ridge axis, whereas a mechanical plate thickness of the order of 6–10 km is derived from regional studies^{30,31}. Lithospheric models of the axial valley¹³ show that the lithosphere acts as an extensional stress guide, episodically and locally broken during accreting periods. Microseismicity and side-scan sonar surveys reveal that the instantaneous position of the spreading axis is marked by volcanic lineaments or tectonic depressions^{15,16}. These are not necessarily located at the centre of the rift valley, but occasionally along a valley wall, indicating that the centre is not necessarily the best place to break the lithosphere. We have calculated the lithospheric stresses induced by small-scale convection in a 2-km-thick elastic plate which is mechanically continuous across the ridge (Fig. 2b). With maximal values of 2 MPa, the tensile

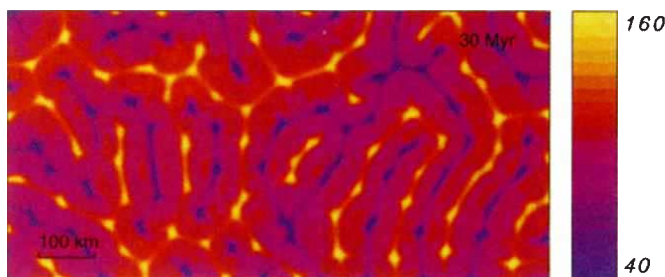


TABLE 1 Physical parameters of model

Acceleration of gravity	g	10 m s^{-2}
Thermal expansion	α	$4.5 \times 10^{-5} \text{ K}^{-1}$
Maximal temperature variation	ΔT	200 K
Layer thickness	d	$6 \times 10^4 \text{ m}$
Mantle density	ρ	$3,300 \text{ kg m}^{-3}$
Viscosity	η	$1.7 \times 10^{18} \text{ Pa s}$
Thermal diffusivity	κ	$7.5 \times 10^{-7} \text{ m}^2 \text{ s}^{-1}$
Elastic thickness	h	2,000 m
Young's modulus	E	$1.5 \times 10^{11} \text{ Pa}$
Poisson's ratio	ν	0.28

stresses in the spreading direction are small compared to the 20 MPa required to maintain the axial valley, but their magnitude is of the same order as the stress drops deduced from microseismicity (0.1–7 MPa)¹¹. As expected, the tensile stresses induced by convection are asymmetric with respect to the axis (Fig. 2b). We infer that they favour the breaking of the lithosphere within the uniformly weakened valley towards the wall

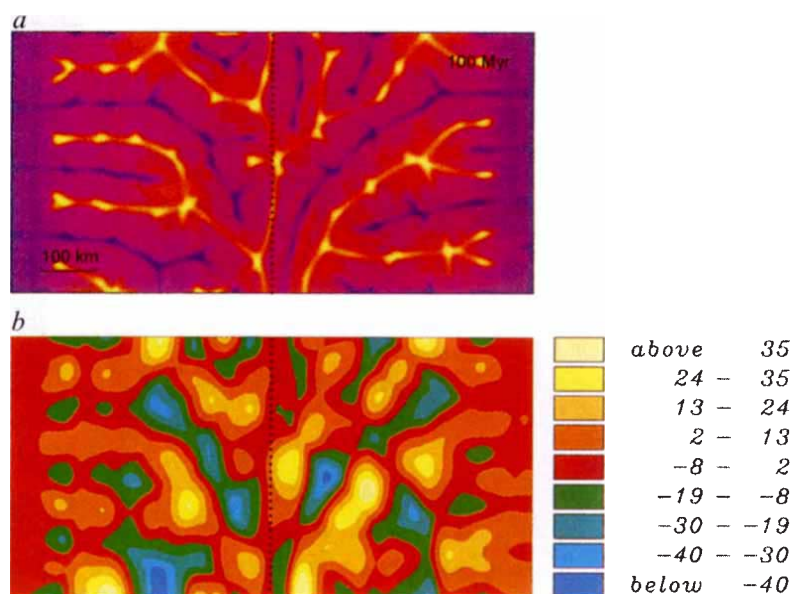
FIG. 1 Horizontal temperature variations at a depth of one-half the height of the model box, after an evolution time of 30 Myr. Units on colour scale, °C. The convective flow occurs in a 60-km-thick, constant-viscosity layer, heated from below and cooled along the fixed upper boundary. The box is $480 \times 960 \text{ km}$. Flow is initiated by small random perturbations of the conductive thermal field. The physical parameters in the model are given in Table 1. With a Rayleigh number of 5×10^4 , the flow pattern is bimodal²⁷ and consists of rolls oriented in various directions. At the intersection of ascending sheets of adjacent rolls, hot channels produce ascending plumes, in which vertical velocities reach 8 cm yr^{-1} . The mean distance between two plumes is 100 km. The flow pattern is far from being steady, but evolves slowly to form rolls parallel to the boundaries of the box. The bimodal steady flow would require 1 Gyr to develop, much longer than the timescale of ridge processes considered here.

Fig. 2 a, Thermal field obtained when the flow pictured in Fig. 1 is symmetrically sheared for 70 Myr by two plates on the top of the model box drifting apart at 1.6 cm yr^{-1} . Colour scale as in Fig. 1. The numerical method used to calculate the forced flow is described elsewhere⁷. Because the vertical velocity of 8 cm yr^{-1} due to buoyancy is much faster than the half-spreading rate of 1.6 cm yr^{-1} , the forced flow is not strong enough to sweep away the whole three-dimensional pattern of buoyant flow²⁸. Ascending currents form vertical sheets on the sides of irregular polygonal cells. These are not symmetrical with respect to the spreading axis, especially below the ridge. This situation remains intrinsically three-dimensional even after 150 Myr (not shown). b, Contours of the normal stress in the spreading direction, induced in an elastic lithosphere of thickness $h = 2 \text{ km}$ by the propagation of the shear stress K_{xz} and K_{yz} generated along the upper interface by the convective flow shown in a. Units on colour scale bar, 0.1 MPa. Elastic stresses are calculated assuming that the shear stress components σ_{xz} and σ_{yz} satisfy $\sigma_{xz} = (z/h)K_{xz}$ and $\sigma_{yz} = (z/h)K_{yz}$ where z is the depth in the plate. The thin plate approximation leads to: $(\sigma_{xx} = (E/(1-\nu^2)) \times [(\partial u/\partial x) + \nu(\partial v/\partial y)]$, where u and v are the horizontal displacements given by:

$$\nabla^2 u = \left(\frac{1-\nu^2}{E} \frac{\partial^2}{\partial x^2} + \frac{2(1+\nu)}{E} \frac{\partial^2}{\partial y^2} \right) K_{xz} - \frac{(1+\nu)^2}{E} \frac{\partial^2}{\partial x \partial y} K_{yz}$$

$$\nabla^2 v = \left(\frac{2(1+\nu)}{E} \frac{\partial^2}{\partial x^2} + \frac{1-\nu^2}{E} \frac{\partial^2}{\partial y^2} \right) K_{yz} - \frac{(1+\nu)^2}{E} \frac{\partial^2}{\partial x \partial y} K_{xz}$$

Equations are solved using the Fourier transforms of K_{xz} and K_{yz} . Propagation of elastic stresses in the whole plate induces two important effects. First, the small magnitude of shearing (0.1 MPa) generated by the convective flow induces stresses in the lithosphere which are an order of magnitude stronger. Second, tensile maxima are not neces-



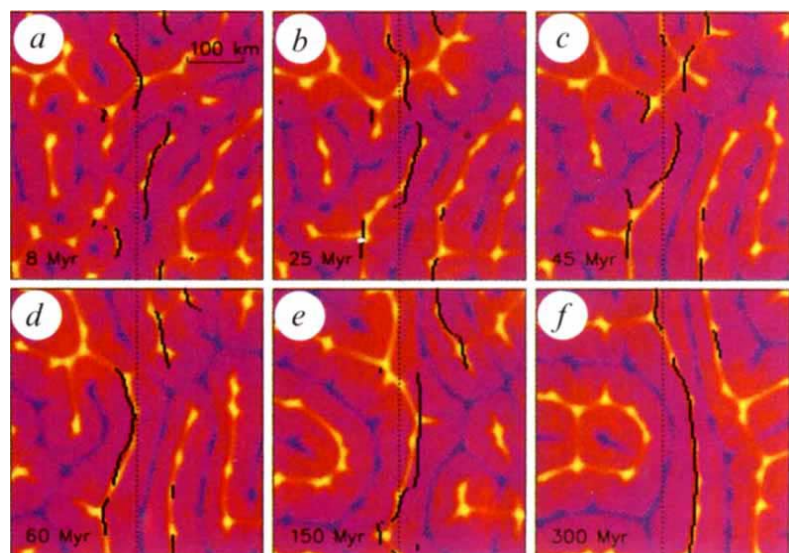
sarily located above ascending currents. As the mantle plumes are not located below the spreading axis, they produce discontinuous local tensile maxima (in yellow) which are clearly asymmetrically located with respect to the ridge axis.

closest to the local maximum for elastic tension. In the following models, each point of the spreading axis is allowed to migrate in the spreading direction by 3.7 km (the half-width of the valley), after each 5 km of plate separation. The incremental shift of the spreading axis modifies the forced flow, inducing in turn a coupling with the buoyant flow. In these conditions, the ridge axis progressively migrates towards the tensile stress maxima near the ascending currents. Because the local tensile maxima form a discontinuous set of lines (Fig. 2b), the ridge develops into separate segments each ~50–150 km long, with offsets ranging from 10 to 100 km (Fig. 3a).

As the convective flow evolves, the ridge geometry continues to change. Figure 3 shows snapshots of the thermal field and ridge positions during a 300-Myr period of evolution. Figure 4a shows the isochrons resulting from the first 100 Myr of ridge evolution. These figures reveal two important observations. First, some offsets shorten with time (Fig. 3a–d and offsets B

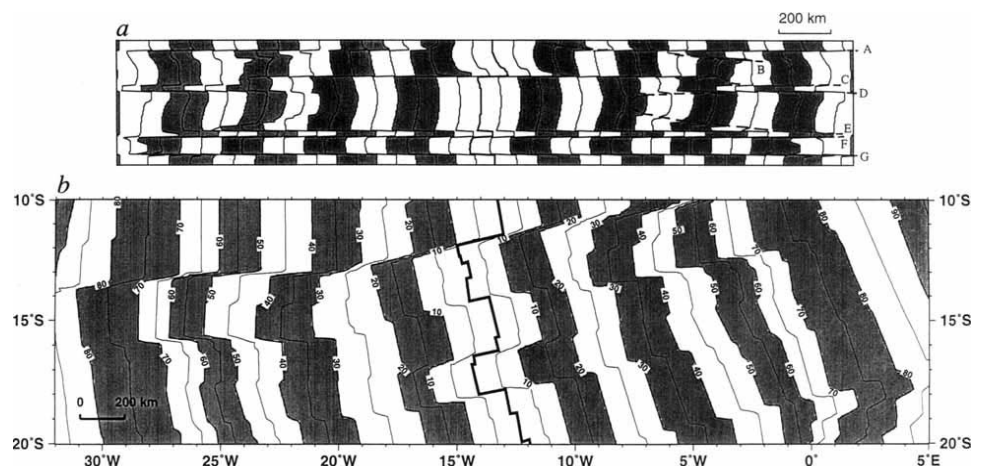
and D in Fig. 4a), as adjacent tensile maxima merge together. Second, the lengths of the segments vary with time (Fig. 3 and segments AB and DE in Fig. 4a). A segment shortens as the tension maximum below it progressively collapses. As a result of this process, the tips of two segments adjacent to an intervening and shortening segment progressively migrate towards the vanishing tensile zone, mimicking along-axis propagation (discontinuities D and E in Fig. 4). As the modelled ridge points are allowed to move only in the spreading direction, this striking result indicates that the models account for the along-axis movements of some mid-ocean ridge segments. Finally, the ridge evolution slows down after 60 Myr, as convective currents form long sheets roughly parallel to the sides of the box (Fig. 3f). The model segmentation resembles that observed on the Mid-Atlantic Ridge (MAR)^{17–20} (Fig. 4), or the Australian–Antarctic discordance^{21–24}, in geometrical pattern, characteristic length and timescale. In particular, offsets of the ridge in the model

FIG. 3 Details of the thermal field obtained when the convective flow shown in Fig. 1 is sheared from above by two diverging plates whose boundary, the ridge axis, moves incrementally. Colour scale as in Fig. 1. Starting as a straight line, the ridge axis is allowed to drift by a 3.7-km step each 0.32 Myr (1.15 cm yr^{-1}), towards the areas of maximum tension. Ages indicate time after the ridge is released. We determine the successive positions of the ridge axis, which are then used to compute the forced flow driven by the plates in their new configuration. Finally, this forced flow is added to the buoyant flow to update the thermal field. In less than 10 Myr, the spreading axis splits into several segments located at tensile maxima (a). This segmentation may seem surprising in light of the continuity of the oblique sheets in the ascending currents below the ridge. Nonetheless, the ridge location is controlled solely by the horizontal σ_{xx} component of the stress field (Fig. 2b). This field displays discontinuous crest lines, explaining the ridge segmentation by five sharp offsets, 10–100 km long, over a distance of <500 km. Some short portions of the ridge (as small as single grid points) temporarily migrate very far from the major segments. These unrealistic segments may be due to the lack of along-axis cohesion of the elastic plate in our modelling. The evolution of the convective flow, however, is found to interact only with the largest segments. During the first 60 Myr, the spreading segments and the ascending flow beneath the ridge system interact to form a quasi-continuous line. Most of the offsets shorten, leaving only the major discontinuities associated with downwelling currents. The segments do not necessarily coincide with the ascending flow: gaps up to



50 km are common close to the offsets. Then the convective pattern continues to evolve, but on a timescale of 100 Myr. In f are shown three ridge segments located above the ascending currents of two distinct transverse rolls.

FIG. 4 a, Stages of the first 90 Myr of the evolution of the same model as in Fig. 3, shown as isochrons every 5 Myr. The ridge axis is the configuration after 90 Myr of evolution (Fig. 3d), while the outermost isochrons represent the geometry of the ridge after 5 Myr of evolution. Some discontinuities (A–G) can be traced from one isochron to another, revealing the shortening of some offsets, and the shortening or lengthening of some segments. b, Ages of the sea floor in the southern Atlantic²⁹, shown with a colour change every 10 Myr, and isochrons every 5 Myr. The spatial scale is the same as in a. The segmentation of the model and its evolution in time strikingly resembles that of the Mid-Atlantic Ridge.



occur in opposite senses, leading to a crenulated pattern very reminiscent of the observations.

The ridge and convective flow evolution described above are clearly strongly dependant on the initial convective field, whose asymmetry initiates the segmentation. Because of the infinite variety of such initial conditions, the geometry and evolution of the ridge segmentation are very different from one case to another. Some characteristics of the ridge segmentation, however, remain in all cases. For instance, the crenulated ridge observed in the experiment shown in Figs 3 and 4a is obtained whenever the initial orientation of the ridge is perpendicular to the spreading direction. Starting with a ridge oblique to the spreading direction, the convective flow induces a segmentation with offsets in the same sense, preserving the overall initial orientation of the ridge. The latter cases account for the geometry observed in the northern MAR²⁵. In all cases, the drift of the ridge segments in the direction of spreading can explain the asymmetric estimates of crustal thickness on the flanks of the MAR from Bouguer gravity residuals²⁶.

Although the conditions required to produce the ridge segmentation depicted here are simple, they do occur in nature. They result from the elastic rheology of a part of the lithosphere in the valley, allowing the tensile stress maxima induced by the three-dimensional convective flow to form a discontinuous set. These stresses are large enough to create perturbations in the large-scale tectonic stresses responsible for the lithospheric necking, leading to shifts in the spreading axis. Asymptotically, all the model runs lead to segments 100–200 km long, orthogonal to the spreading direction. Any change in spreading direction or the interaction between the ridge and the mantle beneath are likely to lead to a reorganization of the convective flow, generating a new segmentation process. Our approach emphasizes the role played by mantle dynamics in triggering ridge segmentation.

Obviously, lithospheric failure and ridge jumps due to unsteady mantle flow may follow more complex rules than the simple but reasonable one adopted here; future studies should lead to an elucidation of what these rules might be. □

Received 6 December 1994; accepted 13 March 1995.

1. Phipps Morgan, J. & Parmentier, E. M. *J. geophys. Res.* **90**, 8603–8612 (1985).
2. Pollard, D. D. & Aydin, A. *J. geophys. Res.* **89**, 10017–10028 (1984).
3. Macdonald, K. C., Scheirer, D. S. & Carbotte, S. M. *Science* **253**, 986–994 (1991).
4. Sandwell, D. T. *J. geophys. Res.* **91**, 6405–6517 (1986).
5. Phipps Morgan, J. *Rev. Geophys. (suppl.)* **29**, 807–822 (1991).
6. Whitehead, J. A., Dick, H. J. B. & Schouten, H. *Nature* **312**, 146–148 (1985).
7. Rabinowicz, M., Rouzo, S., Sempéré, J.-C. & Rosemberg, C. *J. geophys. Res.* **98**, 7851–7869 (1993).
8. Sparks, D. W., Parmentier, E. M. & Phipps Morgan, J. *J. geophys. Res.* **98**, 21977–21996 (1993).
9. Mutter, C. Z. & Mutter, J. C. *Earth planet. Sci. Lett.* **117**, 295–317 (1993).
10. White, R. S., McKenzie, D. & O'Nions, R. K. *J. geophys. Res.* **97**, 19683–19715 (1992).
11. Toomey, D. R., Solomon, S. C. & Purdy, G. M. *J. geophys. Res.* **93**, 9093–9112 (1988).
12. Tappinier, P. & Francheteau, J. *J. geophys. Res.* **83**, 3955–3970 (1978).
13. Chen, Y. & Morgan, W. J. *J. geophys. Res.* **95**, 17571–17581 (1990).
14. Neumann, G. a. & Forsyth, D. W. *J. geophys. Res.* **98**, 17891–17910 (1993).
15. Kong, L. S. L., Detrick, R. S., Fox, P. J., Mayer, L. A. & Ryan, W. B. F. *Mar. geophys. Res.* **10**, 59–90 (1988).
16. Smith, D. K. & Cann, J. R. *Nature* **365**, 707–715 (1993).
17. Carbotte, S., Welch, S. M. & Macdonald, K. C. *Mar. geophys. Res.* **13**, 51–80 (1991).
18. Fox, P. J., Grindlay, N. R. & Macdonald, K. C. *Mar. geophys. Res.* **13**, 1–20 (1991).
19. Grindlay, N. R., Fox, P. J. & Macdonald, K. C. *Mar. geophys. Res.* **13**, 21–50 (1991).
20. Cande, S. C., LaBrecque, J. L. & Haxby, W. F. *J. geophys. Res.* **93**, 13479–13492 (1988).
21. Weissel, J. K. & Hayes, D. E. *J. geophys. Res.* **79**, 2579–2587 (1974).
22. Palmer, J., Sempéré, J.-C., Christie, D. M. & Phipps Morgan, J. *Mar. geophys. Res.* **15**, 121–152 (1993).
23. Weissel, J. K. & Hayes, D. E. *Nature* **213**, 518–521 (1971).
24. Hayes, D. E. *Geol. Soc. Am. Bull.* **87**, 994–1002 (1976).
25. Sempéré, J.-C., Lin, J., Brown, H. S., Schouten, H. & Purdy, G. M. *Mar. geophys. Res.* **15**, 153–200 (1993).
26. Pariso, J. E., Sempéré, J.-C. & Rommevaux, C. *J. geophys. Res.* (submitted).
27. Richter, F. M. & Parsons, B. *J. geophys. Res.* **80**, 2529–2541 (1975).
28. Parmentier, E. M. & Phipps Morgan, J. *Nature* **348**, 325–328 (1990).
29. Müller, R. D., Royer, J.-Y. & Lawver, L. A. *Geology* **21**, 275–278 (1993).
30. Cochran, J. R. *J. geophys. Res.* **84**, 4713–4729 (1979).
31. Blackman, D. K. & Forsyth, D. W. *J. geophys. Res.* **96**, 11741–11758 (1991).

ACKNOWLEDGEMENTS. We thank J.-C. Sempéré and K. Feigl for helpful criticism.

Microwear on conodont elements and macrophagy in the first vertebrates

Mark A. Purnell

Department of Geology, University of Leicester, University Road, Leicester LE1 7RH, UK

FEEDING mechanisms may hold the key to understanding the selective pressures that led to the evolution of vertebrates^{1–3}. But in the absence of direct evidence, hypotheses of feeding mechanisms in the most primitive extinct vertebrates are somewhat speculative, and scenarios that seek to explain vertebrate origins remain controversial. The recognition that the affinities of conodonts lie among the most primitive vertebrates^{4–6} shifts the balance in this debate. I illustrate here microscopic wear patterns on conodont elements that provide the first unequivocal evidence that they functioned as teeth. These microwear patterns were produced as food was crushed and sheared between opposed conodont elements brought into bilateral occlusion. The presence of teeth and evidence of macrophagy in such primitive and early vertebrates support hypotheses that the first vertebrates were predators.

Although recent hypotheses have highlighted the central roles of feeding mechanisms and mode of life in understanding the origin of vertebrates^{1–3}, there is no direct evidence to indicate how early agnathan fishes fed. This has resulted in a lack of consensus regarding feeding in the first vertebrates. The traditional view contends that they were microphagous pump-suspension feeders (for example, refs 3, 7), a hypothesis based on comparisons with amphioxus and larval lampreys, and fossil

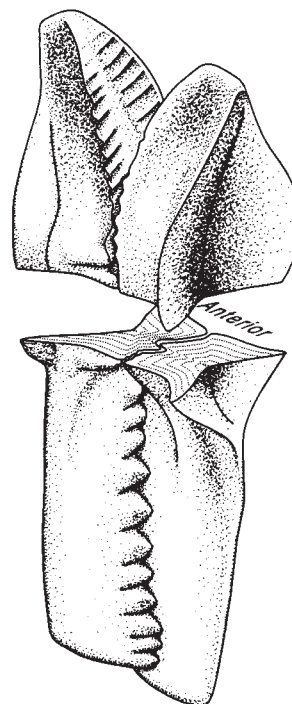


FIG. 1 Bilateral occlusion in ozarkodinid conodonts. The ornamented platform areas of elements were brought into contact by rotation about a hinge point located close to the junction between the platform and the blade. Blade-shaped ozarkodinid elements (not illustrated) also rotated in a similar manner⁹. Illustration is diagrammatic, but based on paired platform (Pa) elements of *Idiognathodus*.

evidence that early agnathans had heavy dermal armour, lacked adaptations for fast swimming and lacked biting structures. The alternative hypothesis proposes that many of the characters that define vertebrates, such as paired sensory organs and muscular and skeletal adaptations for active life, would not have evolved unless the earliest vertebrates were predatory^{1,2}.

Until recently, conodonts had no relevance to this debate, but it is now recognized that their anatomy and skeletal histology indicate that they were primitive vertebrates^{4,5,8} (although see ref. 4 for a discussion of alternative interpretations). Conodonts first appear in Upper Cambrian strata, 40 million years earlier than other vertebrate remains and, in marked contrast to the paucity of evidence for feeding in early agnathans, the conodont fossil record is made up almost exclusively of their phosphatic mouth parts. However, hypotheses of feeding in conodonts are

controversial. Recent analyses of functional morphology and ontogeny have suggested that conodont elements functioned as teeth^{9,10}, but this hypothesis fails a critical test. The surface of conodont elements is composed of lamellar apatite similar in structure to¹¹, and possibly homologous with, enamel⁵; if they functioned as teeth, therefore, conodont elements should exhibit wear patterns similar to those documented in the teeth of higher vertebrates. The observation that they do not (for example, refs 12–14) is difficult to reconcile with hypotheses of tooth-like function.

Wear has been noted in the teeth of a number of vertebrate groups but systematic study of the microscopic features of dental wear has been restricted to mammals and advanced mammal-like reptiles¹⁵. This 'microwear' is produced *in vivo* by abrasives in food and by the compressive and shearing forces that act on

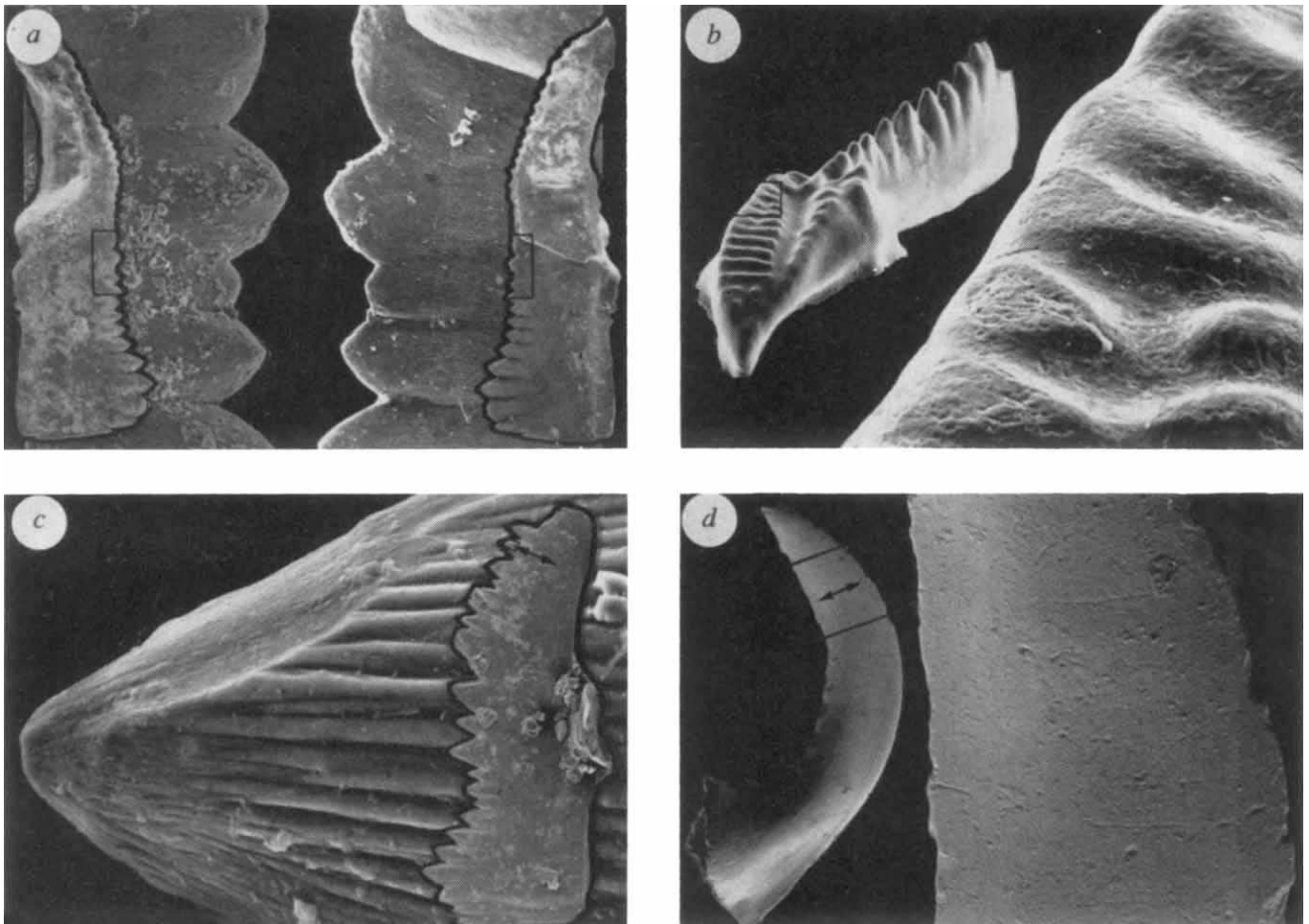


FIG. 2 Microwear in conodont elements. a, Pa element of *Gnathodus bilineatus*; the original fibrous surface texture is well preserved on the forward facing surface of the element (shown on left) but has been worn off the backward facing surface (shown on right) where it was in contact with the opposed element. This wear, on one side only, cannot be the result of post-mortem abrasion. The worn areas are smoothly polished with no microfeatures discernible; such smooth polishing or frosting in teeth indicates tooth on tooth contact, either in the absence of food¹⁸ or with non-abrasive food²³. Carboniferous, Stonehead Beck, North Yorkshire; Leicester University specimen LEUIG 114245, whole element $\times 29$, close up $\times 232$. b, Pa element of *Idiognathodus*; the crests of the platform ridges in this element are blunted and flattened to form triangular wear facets with pitted microfeatures. Such features are not developed elsewhere on the element and are most unlikely to be post-mortem artefacts. Pitting or chipping of facets is diagnostic of crushing or compression^{16,24}. Upper Carboniferous, Elk County, Kansas,

USA; Royal Ontario Museum specimen ROM 50699, whole element $\sim \times 36$, close up $\sim \times 387$. c, Pa element of *Ozarkodina confluens*; a well developed wear facet formed by contact with the opposed element sharply truncates the original fibrous surface texture. The facet is covered with fine, parallel striations which indicate that this element was used for shearing^{17,18,24}; arrows indicate orientation of shearing motion. The distinctive nature of this facet and its location, between denticles, preclude a post-mortem origin. Silurian, Prior's Frome, Hereford and Worcester; Natural History Museum Specimen X-1299, whole element $\times 39$, close up $\times 1,367$. d, Element of *Drepanoistodus*, apex tilted a little towards viewer; a number of straight, parallel scratches traverse the slightly convex surface; their parallelism precludes a post-mortem origin. This microwear is characteristic of surfaces used predominantly for shearing^{17,18,24}; arrows indicate orientation of shearing motion. Ordovician, Saudi Arabia, Natural History Museum Specimen X-1300, whole element $\sim \times 36$, close up $\sim \times 194$.

enamel during feeding¹⁵⁻¹⁷. Thus, it records information about the mechanics of food breakdown, the relative motion of teeth, and the nature of food consumed. Analysis of microwear therefore represents a very powerful and direct tool for investigating feeding in fossil animals, but its application to fossil material is complicated by the problem of post-mortem abrasion.

In studies of fossil mammal teeth, post-mortem abrasion can be readily identified because *in vivo* wear forms as distinctive facets at specific locations on functional surfaces¹⁸. In most cases, these functional surfaces were occlusal and, although complex occlusion is generally considered characteristic of mammals¹⁹, a similar condition arose in conodonts^{14,20,21} more than 200 million years earlier. Occlusion in conodonts, however, was bilateral, and in the order Ozarkodinida the left element of a pair fitted behind that on the right (Fig. 1; and manuscript in preparation). This consistent left behind right pairing indicates that rather than being related to asymmetrical feeding behaviour²², morphological asymmetry of ozarkodinid elements represents genetically controlled developmental lateralization to allow occlusion. More importantly, however, recognition of left behind right occlusion enables the identification of functional surfaces of conodont elements. Detailed scanning electron microscopy (SEM) examination of these surfaces has revealed the presence of wear that cannot be post-mortem in origin (Fig. 2) and must have been produced *in vivo*.

Three basic types of microwear occur on mammal teeth: frosting or smooth polishing, fine scratching or striation, and pitting or chipping. The same microwear patterns occur on conodont elements and, in contrast to previous general and sometimes speculative statements of possible 'tooth-like' function, they allow precise characterization of food processing in conodonts. Thus, the polished areas on the blade of *Gnathodus bilineatus* (Fig. 2a) indicate that this part of the element was not in contact with food, or that the species ate food that was not abrasive (compare with refs 18, 23). The pitted microwear on *Idiogonathodus* platforms (Fig. 2b) indicates that food was crushed between opposed elements, but the lack of associated scratches suggests that they did not grind^{16,24}. The parallel scratching on elements of *Ozarkodina confluens* (Fig. 2c) and *Drepanoistodus* (Fig. 2d) is diagnostic of shearing^{17,18,24}. The scratching is related to the abrasiveness of the food consumed¹⁷ and it is probable that *Drepanoistodus* ate food which contained more abrasive particles than that of *O. confluens*. Also, the degree of parallelism of scratches on teeth reflects the tightness of occlusal guidance²⁵; given that conodonts were jawless, the scratching illustrated suggests surprisingly precise control of element contact. This is especially true of *Drepanoistodus* as this genus bore only simple cone-shaped elements and belongs to a primitive order of conodont dating back to the Late Cambrian²⁶.

The presence of wear in conodonts has a number of implications for models of element growth. Wear has been observed on small, immature elements, indicating that conodonts did not grow their teeth to full size before using them. Also, internal discontinuities in conodont growth that have in the past been interpreted as resorption²⁷ may be early phases of surface wear followed by growth. If so, this cycle of alternating growth and use falsifies ontogenetic hypotheses of element shedding and replacement^{28,29}, but makes element histogenesis and growth somewhat problematic. More generally, the interpretation of feeding and mode of life in conodonts and extinct agnathans has been rather speculative, but microwear provides a means of direct analysis and hypothesis testing. For example, some osteostracans³, heterostracans³ and conodonts³⁰ may have been deposit feeders; the feeding elements of such animals should exhibit microwear features that reflect the abrasiveness of their food.

Microwear analysis of conodont function is at a preliminary stage, but the qualitative results presented here, the first application of microwear analysis to non-tetrapods, give an indication of the potential of the methodology. The microwear features

illustrated provide the first unequivocal evidence that conodont elements were teeth and the occurrence of shearing, a method of food breakdown incompatible with microphagy, indicates that conodonts were macrophagous. Conodonts also had good vision³¹ and were probably capable of rapid anguilliform locomotion^{4,30}. The presence of this suite of characters in a group of such antiquity and primitiveness lends strong support to hypotheses that the earliest vertebrates were active predators, their origins linked with an ecological shift from suspension feeding to predation^{1,2}. □

Received 15 September 1994; accepted 15 February 1995.

- Jollie, M. *Zool. J. Linn. Soc.* **75**, 167-188 (1982).
- Gans, C. & Northcutt, R. G. *Science* **220**, 268-274 (1983).
- Mallatt, J. *Zool. J. Linn. Soc.* **82**, 261-272 (1984).
- Aldridge, R. J., Briggs, D. E. G., Smith, M. P., Clarkson, E. N. K. & Clark, N. D. L. *Phil. Trans. R. Soc. B* **340**, 405-421 (1993).
- Sansom, I. J., Smith, M. P., Armstrong, H. A. & Smith, M. M. *Science* **256**, 1308-1311 (1992).
- Briggs, D. E. G. *Science* **256**, 1285-1286 (1992).
- Romer, A. S. *The Vertebrate Body* 1-601 (Saunders, Philadelphia, 1970).
- Sansom, I. J., Smith, M. P. & Smith, M. M. *Nature* **368**, 591 (1994).
- Purnell, M. A. & von Bitter, P. H. *Nature* **359**, 629-631 (1992).
- Purnell, M. A. *Geology* **21**, 375-377 (1993).
- Schmidt, H. & Müller, K. J. *Paläontol. Z.* **38**, 105-135 (1964).
- Hass, W. H. J. *Paleont.* **15**, 71-81 (1941).
- Pierce, R. W. & Langenheim, R. L. *Bull. Geol. Soc. Am.* **81**, 3225-3236 (1970).
- Nicoll, R. S. in *Palaeobiology of Conodonts* (ed. Aldridge, R. J.) 77-90 (Ellis Horwood, Chichester, 1987).
- Teaford, M. F. in *Advances in Dental Anthropology* (eds Kelley, M. A. & Larsen, C. S.) 341-356 (Liss, New York, 1991).
- Maas, M. C. *Arch. Oral Biol.* **39**, 1-11 (1994).
- Maas, M. C. *Am. J. Phys. Anthropol.* **85**, 31-49 (1991).
- Teaford, M. F. *Scan. Microsc.* **2**, 1167-1175 (1988).
- Janis, C. M. in *Evolutionary Paleobiology of Behavior and Coevolution* (ed. Boucot, A. J.) 241-259 (Elsevier, Amsterdam, 1990).
- Jeppsson, L. *Lethaia* **4**, 101-123 (1971).
- Weddige, K. *Cour. Forsch. Inst. Senckenberg* **118**, 563-589 (1990).
- Babcock, L. E. J. *Paleont.* **67**, 217-229 (1993).
- Rensberger, J. M. in *Development, Function and Evolution of Teeth* (eds Butler, P. M. & Joysey, K. A.) 415-438 (Academic, London, 1978).
- Gordon, K. D. *Am. J. Phys. Anthropol.* **59**, 195-215 (1982).
- Gordon, K. D. *Scan. Microsc.* **2**, 1139-1147 (1988).
- Aldridge, R. J. & Smith, M. P. in *The Fossil Record 2* (ed. Benton, M. J.) 563-572 (Chapman and Hall, London, 1993).
- Müller, K. J. in *Treatise on Invertebrate Paleontology, Part W, Miscellaneous, Supplement 2, Conodonts* (ed. Robison, R. A.) W20-W41 (Geological Society of America and University of Kansas, Lawrence, 1981).
- Carls, P. *Neues jb. geol. paläont. Abh.* **155**, 18-64 (1977).
- Krejsa, R. J., Bringsa, P. & Slavkin, H. C. *Lethaia* **23**, 359-378 (1990).
- Aldridge, R. J. & Briggs, D. E. G. in *Problematic Fossil Taxa. Oxford Monographs on Geology and Geophysics No. 5* (eds Hoffman, A. & Nitecki, M. H.) 227-239 (Oxford Univ. Press, New York, 1986).
- Purnell, M. A. *Lethaia* (in the press).

ACKNOWLEDGEMENTS. I thank R. J. Aldridge, P. H. von Bitter and D. Martill for discussion, access to material and comment on the manuscript. R. Branson and G. McTurk provided photographic and microscopic assistance. Funded by a NERC Fellowship.

A giant conodont with preserved muscle tissue from the Upper Ordovician of South Africa

S. E. Gabbott*, R. J. Aldridge* & J. N. Theron†

* Department of Geology, Leicester University, Leicester LE1 7RH, UK
† Geological Survey, PO Box 572, Bellville 7535, South Africa

AN exceptionally preserved new specimen of the giant conodont *Promissum pulchrum* reveals details of the trunk musculature, feeding apparatus and eyes. High-fidelity resolution of ultrastructural features of the trunk myomeres provides the first conclusive evidence of muscle-fibre organization and orientation in an extinct agnathan. The presence of fibrous extrinsic eye muscles confirms the degree of encephalization of the conodonts and is consistent with a cladistic position crownwards of the myxinoids. The soft tissues are uniquely preserved as illite and mixed-layer clay minerals.

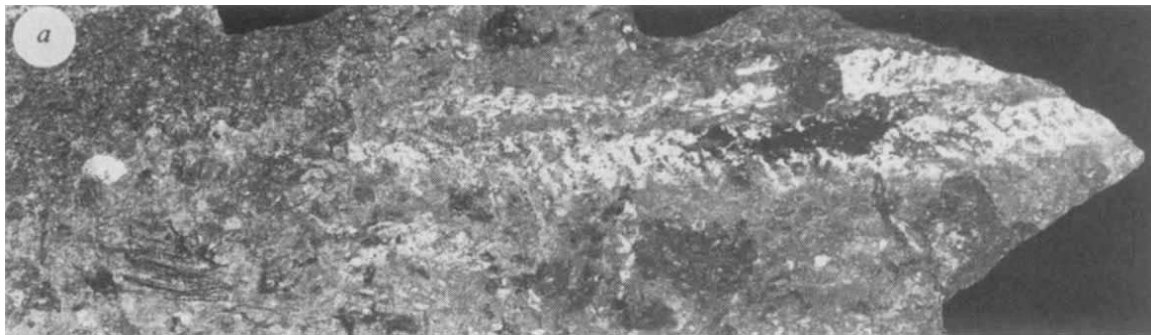
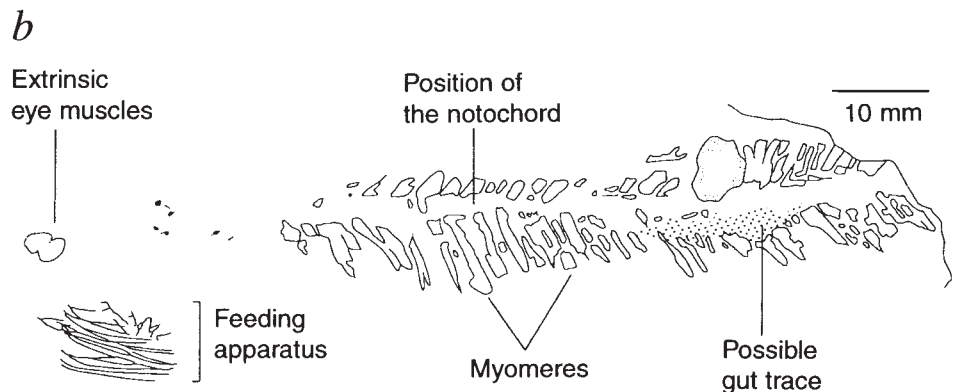


FIG. 1 *Promissum pulchrum* Kovács-Endrődy, part (specimen C721a, Geological Survey of South Africa, Pretoria). a, Photograph $\times 1.3$. b, Interpretative camera-lucida drawing.



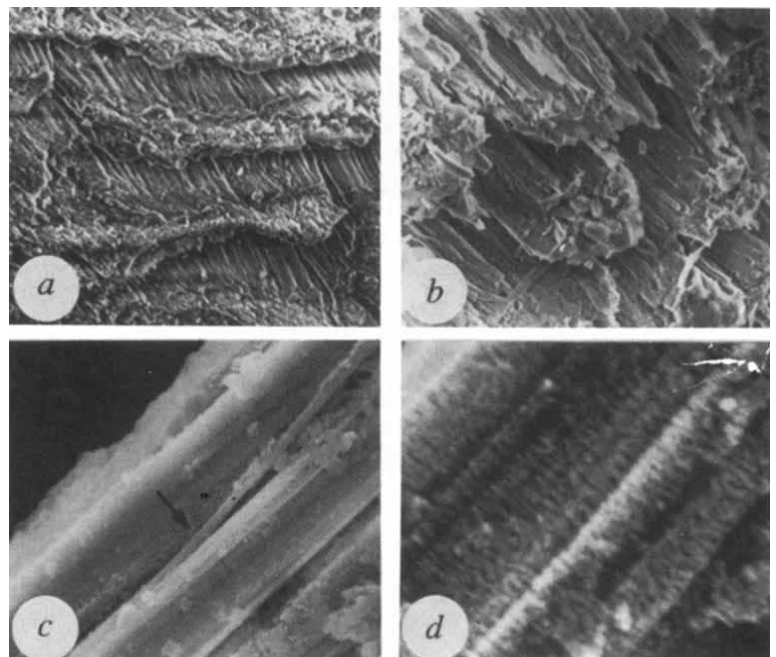
Conodonts are extinct Cambrian/Triassic vertebrates, soft bodied apart from the phosphatic elements of their feeding apparatuses¹. Specimens preserving the soft tissues are very rare; ten almost completely preserved animals are known from the Carboniferous Granton Shrimp Bed of Scotland¹⁻³ and one faint partial specimen from the Silurian Brandon Bridge Dolomite of Wisconsin^{4,5}. Traces of sclerotic capsules surrounding the eyes have previously been reported in association with exceptionally large apparatuses of *Promissum pulchrum* in the Upper Ordovician Soom Shale of South Africa⁶, where the more complete remains of a single animal have been discovered.

The Soom Shale is a black, finely laminated argillaceous unit

within the dominantly arenaceous Lower Palaeozoic Table Mountain Group of south-west South Africa^{7,8}. The shale overlies tillites and was deposited during the Late Ordovician glaciation of Gondwanaland. Several exceptionally preserved fossils have been found in the Soom Shale⁸, making it the most important Ordovician Konservat-Lagerstätte and the only such deposit found in cold, high-latitude (60° S), glacially influenced marine waters.

The conodont specimen comes from the farm of Sandfontein (19°14' E, 32°40' S), 150 km north of Cape Town. Only the anterior portion of the animal was found, represented by part and counterpart. The preserved length is 109 mm (Fig. 1). Linear

FIG. 2 Scanning electron micrographs of muscle fibres within a myomere of *Promissum pulchrum*. a, Layers of rod-like fibres plunging away from myoseptum, $\times 144$. b, Fragment of muscle tissue showing fibrous texture, $\times 530$. c, Fibres approximately 5 μm across, showing longitudinal lineation reflecting myofibrillar structure and detachment of a myofibril (arrowed), $\times 2,250$. d, Detail of a single fibre showing possible sarcomeres of striated muscle, $\times 12,000$.



scaling against a complete Carboniferous animal (specimen IGSE 13821 (ref. 2); S-element length 0.76 mm, trunk height 1.8 mm, total length 40.5 mm) suggests that the *Promissum* animal (S-element length 10 mm, trunk height 18 mm) would have approximated 400 mm in its entire length. The configuration of the body and the feeding apparatus attests to lateral collapse and subsequent flattening of the animal. The eyes are positioned above and anterior to the apparatus, which comprises 19 denticulated elements and has been interpreted as a grasping and crushing array⁹. The most anteriorly preserved V-shaped myomere occurs 29 mm behind the eyes; that the chevrons are discrete is attributable to post-mortem shrinkage of the muscle fibres before fossilization¹⁰. A longitudinal unmineralized band 2 mm high is interpreted by comparison with the Carboniferous animals to mark the position of the notochord (Fig. 1); there is no trace of the notochord itself on the part or counterpart of the Soom specimen. Ventral to this, 63 to 84 mm posterior of the eyes, an elongate black patch of structureless organic material within the body of the animal may be a relic of the gut or other visceral organ.

Promissum belongs to the conodont order Prioniodontida, with an architecturally different feeding apparatus from the Scottish Carboniferous specimens, which belong to the Ozarkodontida⁹. The preserved soft-tissue anatomy of *Promissum*, however, compares closely with that of the Granton animals, although the sclerotic capsules representing the eyes are smaller^{1,11}. Possible fibrous muscle tissue is evident in the myomeres of one of the Granton specimens¹, but the fidelity of replication of muscle tissue is greater in the Ordovician specimen.

Each myomere of the *Promissum* animal shows a fibrous texture which plunges obliquely into the sediment (Fig. 2a, b). In the horizontal plane, the dominant orientation of the fibres is at 10° to the long axis of the animal, rising anteriorly; this is possibly the result of post-mortem shrinkage shifting the fibres from an original longitudinal lineation. Scanning electron microscopy of the fibres reveals similarities to muscle tissue preserved in

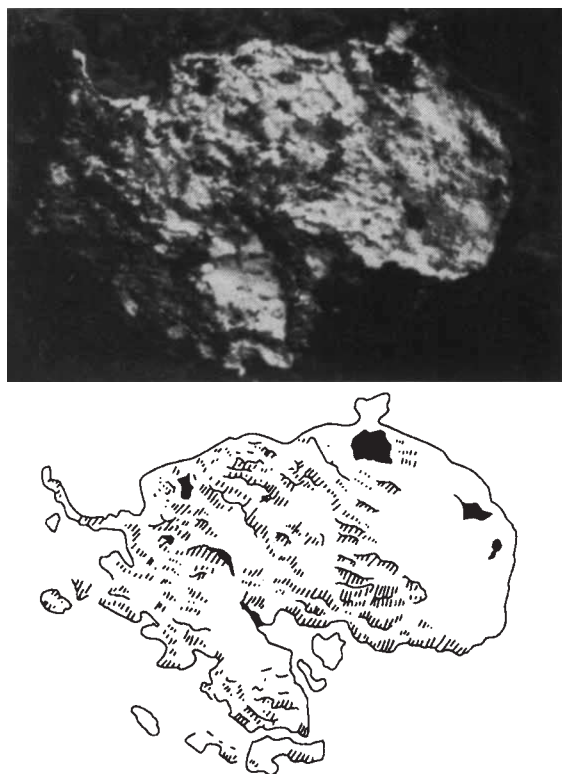


FIG. 3 Photograph and camera-lucida drawing of the extrinsic eye muscles, showing fibrous texture, $\times 15$.

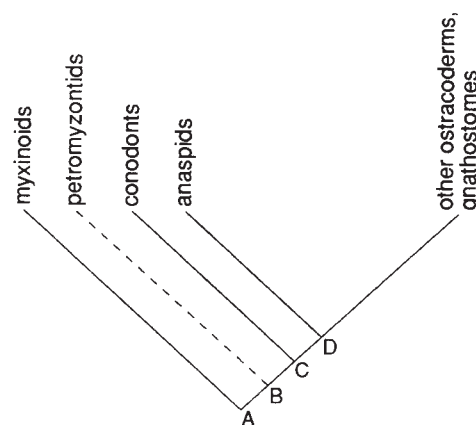


FIG. 4 Simplified and modified hypothesis of conodont relationships (after refs 1, 18). Characters at node A include neural crest, cranium, bilaterally operative feeding apparatus, paired olfactory, optic and auditory capsules; at node B, extrinsic eye muscles (if not at A); at node C, vertebrate hard tissues; at node D, paired fins, muscularized unpaired fins. If the mineralized skeleton and paired fins are secondarily absent in petromyzontids, they move crownwards of the conodonts, perhaps as a sister group of the anaspids.

fossil fish¹² and in laboratory experiments mimicking the phosphatization of arthropod musculature during fossilization^{13,14}. Fibres with fibrils, sarcolemmic membranes and possible collagenous connective tissues are evident. Lineation in the muscle fibres reflects fibril bundles which may split away (Fig. 2c, arrowed). In some places a poorly developed transverse lineation is comparable in style with that shown by fossilized striated muscle¹², but is finer in scale (Fig. 2d). We interpret the solid oval patches with fibrous texture anterior to the feeding apparatus as extrinsic eye musculature (Fig. 3); sclerotic eye cartilages of comparable shape and size have been described in the same position relative to the apparatuses on other *Promissum* specimens⁶.

The muscle fibres in the myomeres commonly show a microgranular surface, reflecting the texture of the replacing minerals, and comparable with naturally and experimentally phosphatized muscle tissue^{13,14}. Energy dispersive X-ray and electron microprobe analyses, however, show that the *Promissum* muscles are preserved as illites and mixed-layer clays, not calcium phosphate. This mode of preservation is problematic; it may be that authigenic clays, precipitated at elevated diagenetic temperatures and decreased pH, replaced a primary phosphatic phase of soft-tissue replication. Alternatively, a single phase of replacement may have occurred in which the organic muscle membranes formed a template for direct crystallization of authigenic clays.

This discovery represents the earliest fibrous muscle tissue preserved in the vertebrate fossil record. The extrinsic eye musculature indicates a degree of encephalization for the conodonts comparable with that of the petromyzontids. Myxinooids have poorly developed lateral eyes and lack extrinsic eye muscles, a characteristic that may be interpreted as primitive or degenerate¹⁵. In either case, the eye musculature provides a further character consistent with the phylogenetic position of conodonts as vertebrates. The presence of characteristic vertebrate hard tissues^{16,17} places them crownwards of the myxinooids¹, and of the petromyzontids if these never had a mineralized skeleton¹⁸ (Fig. 4).

The simple V-shaped myomeres of the conodont are similar to those of the cephalochordate *Branchiostoma*, but the fibres do not show the characteristic extreme flattening¹⁵. In fish, including Agnatha, the fibre cross-section is less circular than in *Promissum*, and there are at least two fibre types: small-diameter fibres are slow (red), large-diameter fibres fast (white)¹⁹. The 5- μ m diameter in *Promissum* would be consistent with slow fibres in

an agnathan. If further studies of the muscles fail to reveal larger fibres, this would indicate that *Promissum* was an efficient cruiser but was incapable of bursts of rapid swimming. □

Received 8 February; accepted 22 March 1995.

1. Aldridge, R. J., Briggs, D. E. G., Smith, M. P., Clarkson, E. N. K. & Clark, N. D. L. *Phil. Trans. R. Soc. B* **340**, 405–421 (1993).
2. Briggs, D. E. G., Clarkson, E. N. K. & Aldridge, R. J. *Lethaia* **16**, 1–14 (1983).
3. Aldridge, R. J., Briggs, D. E. G., Clarkson, E. N. K. & Smith, M. P. *Lethaia* **19**, 279–291 (1986).
4. Mikulic, D. G., Briggs, D. E. G. & Kluesendorf, J. *Science* **228**, 715–717 (1985).
5. Smith, M. P., Briggs, D. E. G. & Aldridge, R. J. in *Palaeobiology of Conodonts* (ed. Aldridge, R. J.) 91–104 (Ellis Horwood, Chichester, 1987).
6. Aldridge, R. J. & Theron, J. N. *J. Micropalaeontology* **12**, 113–117 (1993).
7. Theron, J. N., Rickards, R. B. & Aldridge, R. J. *Palaeontology* **33**, 577–594 (1990).
8. Aldridge, R. J., Theron, J. N. & Gabbott, S. E. *Geol. Today* **10**, 218–221 (1994).
9. Aldridge, R. J., Purnell, M. A., Gabbott, S. E. & Theron, J. N. *Phil. Trans. R. Soc. B* **347**, 275–291 (1995).
10. Briggs, D. E. G. & Kear, A. J. *Lethaia* **26**, 275–287 (1994).
11. Purnell, M. A. *Lethaia* (in the press).
12. Martill, D. M. *Nature* **346**, 171–172 (1990).
13. Briggs, D. E. G., Kear, A. J., Martill, D. M. & Wilby, P. R. *J. geol. Soc. Lond.* **150**, 1035–1038 (1993).
14. Briggs, D. E. G. & Kear, A. J. *Palaios* **9**, 431–456 (1994).
15. Northcutt, R. G. in *Evolutionary biology of primitive fishes* (eds Foreman, R. E., Gorbman, A., Dodd, J. M. & Olsson, R.) 81–112 (Plenum, New York, 1985).
16. Sansom, I. J., Smith, M. P., Armstrong, H. A. & Smith, M. M. *Science* **256**, 1308–1311 (1992).
17. Sansom, I. J., Smith, M. P. & Smith, M. M. *Nature* **368**, 591 (1994).
18. Forey, P. & Janvier, P. *Am. Scient.* **82**, 554–565 (1994).
19. Bone, Q. in *Fish physiology* Vol. VII *Locomotion* (eds Hoar, W. S. & Randall, D. J.) 361–424 (Academic, New York, 1978).

ACKNOWLEDGEMENTS. We thank J. D. Kotze for permission to collect at Sandfontein; T. Williams, R. Branson, G. McTurk, A. Smith and R. N. Wilson for help with analyses; C. Brooks for photographic assistance; Q. Bone, D. E. G. Briggs, P. Janvier, D. M. Martill, M. A. Purnell and P. R. Wilby for discussion and comments. Financial support is from NERC (R. J. A.); S.E.G. is a NERC Research Student.

Plant invasion of newly exposed hypersaline Dead Sea shores

Dan Yakir* & Yoseph Yechieli†

* Department of Environmental Sciences and Energy Research, The Weizmann Institute of Science, Rehovot 76100, Israel

† Geological Survey of Israel, Jerusalem 95501, Israel

THE level of the Dead Sea, the lowest (about –400 m) and one of the most salty (salinity about 340 g l⁻¹) lakes on earth, is lowering at a rate of approximately 0.5 m annually owing to extensive exploitation of its main perennial tributary (the Jordan River) and the extreme aridity of the region (annual precipitation is about 60 mm)¹. Consequently, new hypersaline sea shores are exposed, forming a unique, originally sterile ecosystem. The first plants invade these newly exposed shores after several years while soil water salinity is still extremely high. Here we use stable isotopes of oxygen and hydrogen to show that a variety of such perennial pioneer plants are able to make use of occasional floodwater which is distinct from the bulk of the hypersaline soil water found in their root zone. Our results provide new insight into the ways in which plants can invade extremely hostile environments and extend their ecological limits of distribution.

Vegetation at the study site (Fig. 1) was sparse, and the existence of small stumps reflected unsuccessful attempts by plants to establish themselves in the area. The well established plants that were sampled in the restricted study area were individuals of *Sueda fruticosa*, *Anabasis articulata*, *Atriplex halimus*, *Hamada scorparia* and *Tamarix nilotica*, all of which are typical to the region². The salt concentration of soil water in the study area was about one third of the salt concentration of the Dead Sea itself (Table 1), consistent with the exposure time of 10–15 years and with the associated leaching by precipitation and occasional floods from higher elevations³. Notably, however, these ‘diluted’

TABLE 1 Composition of interstitial solution from soil samples from root zones of plants growing on newly exposed Dead Sea shores

Sample	Concentration of element (g l ⁻¹)						
	Na	K	Ca	Mg	Cl	SO ₄	Br
DY12	9.9	3.2	4.8	12.0	62.4	3.0	1.5
DY14	8.1	1.8	5.3	9.6	52.1	2.5	1.3
DY16	14.5	4.0	6.5	16.5	87.6	2.3	2.1
DY198	7.8	2.6	3.8	9.5	48.9	3.1	1.2
Mean (A)	10.1	2.9	5.1	11.9	62.8	2.7	1.5
Dead Sea (B)	40.1	7.6	17.2	44.0	224.9	0.5	5.3
A/B	0.3	0.4	0.3	0.3	0.3	5.4	0.3

The average chemical composition of interstitial solutions is compared to that of the Dead Sea water near the study area (Fig. 1). Soil samples were dug out at different depths after the main root of the plant was traced to a depth where minor, living rootlets could be identified (up to 1 m below the surface). Additional soil samples were analysed for Cl only, with no significant differences to the values reported here.

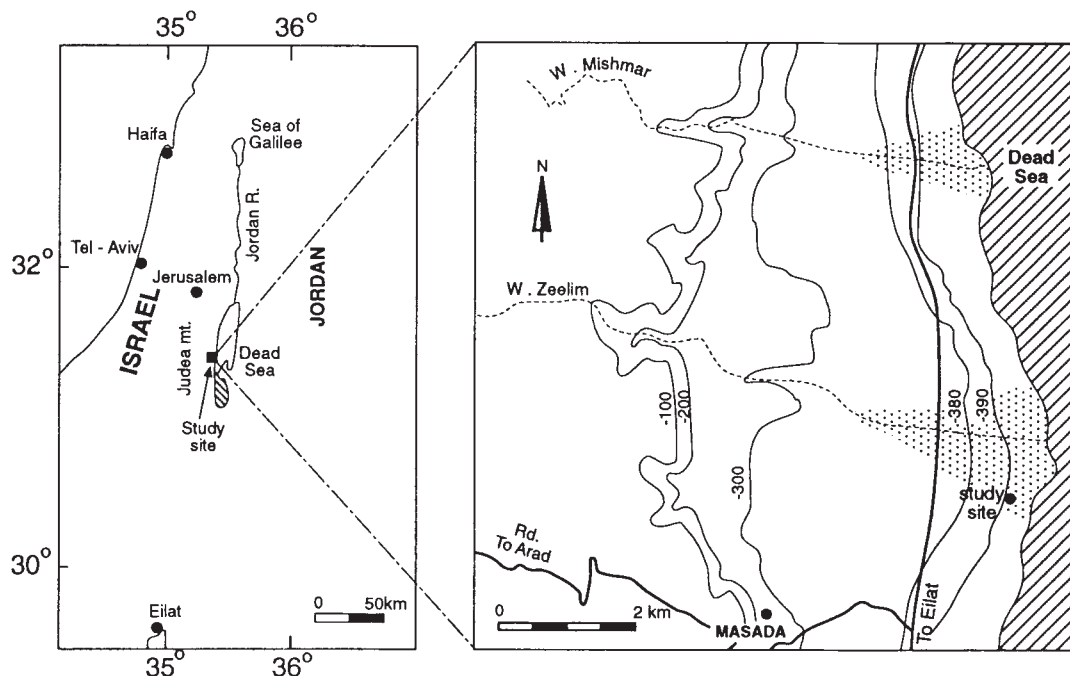
salt concentrations are 2–3 times greater than in ocean water (~35 g l⁻¹), and are well above the levels usually reported to be sustained even by halophytes^{4–6}. We are aware of only one report of a halophyte, from the Dead Sea area, found in soil containing about 8% salts⁷, compared to 9% (by weight) in the present study. It is important to emphasize, however, that the salinity of the soil water does not provide unequivocal information regarding the composition of the actual water used by the plants, and it is in this respect that stable isotopes are invaluable^{8–13}.

As there is no change in the isotope composition of water during uptake by the roots and during flow in the stems, stem water can serve as an indicator of the isotope composition of the actual soil water used by the plants¹³ (but see also ref. 14). In our study area, soil water is originally derived from the Dead Sea which has a very distinct isotopic signature^{15–17} ($\delta^{18}\text{O} \approx +4\text{‰}$; $\delta^2\text{H} \approx +9\text{‰}$ ¹⁸; where $\delta\text{‰} = [(R_{\text{sample}}/R_{\text{standard}}) - 1] \times 10^3$, R is $^{18}\text{O}/^{16}\text{O}$ or $^2\text{H}/^1\text{H}$ and the standard is SMOW (standard mean ocean water)), which is very different from regional precipitation and ground water ($\delta^{18}\text{O} \approx -5\text{‰}$; $\delta^2\text{H} \approx -35\text{‰}$)¹⁹. Notably, irrespective of the isotope composition of the input water, leaf water normally attains higher concentrations of ^{18}O and ^2H owing to preferential loss of water containing the lighter isotopes during evaporation in transpiration.

Plants and soil water were sampled during 1993: the first sampling (limited plant material for ^{18}O analysis only) was during the summer (May) and the second (plant and soil samples for both ^{18}O and ^2H analysis) at the beginning of winter (November) after the first report of rain in the region. Soil samples were dug out from various depths (up to 1 m) while carefully exposing the root system of the plants and identifying living rootlets (indicating an apparently active water-uptake zone). Samples of the roots, stems and leaves were collected and sealed in test tubes. It was also important to obtain a representative sample of flood water, a major source of water in the study area, which could have been isotopically different from local rain. Rocky outcrop and the surface formation of the shallow soil in the Judean desert result, when rainy, in runoff that concentrates and floods the dry riverbeds leading towards the Dead Sea. Flood water, difficult to sample because of its unpredictability and episodic nature, was obtained from a nearby riverbed during the winter of 1991. Its $\delta^{18}\text{O}$ value was within the range of the long-term values of flood water for the region¹⁹ and is reported here as a representative value.

We compared the isotope composition of stem, root and leaf water of the plant species from the study area to the isotope composition of soil water in the active root zone, and to that of the Dead Sea and flood water (Fig. 2). In the plot of $\delta^{18}\text{O}$ against $\delta^2\text{H}$ values for the water samples (Fig. 2), two distinct lines were apparent. The first line connects the Dead Sea water with the soil water (slope 5.2), and the second runs between the

FIG. 1 Map of the Dead Sea region indicating the study area on shores exposed 10–15 years BP. The main dry river beds that carry occasional winter floods are indicated.



leaf water and root water (slope 3.3). This latter line represents a group of individual evaporation lines running between specific root water and the corresponding leaf water with slopes between 2.8 and 3.5, which are typical of evapo-transpiration lines^{20,21}. As expected, the soil water represented a mixture of flood water and Dead Sea water, and the position of average soil water in the $\delta^{18}\text{O}$ versus $\delta^2\text{H}$ plot is consistent with the extent of dilution of the Dead Sea water at the site (Table 1). Seasonal variations in the $\delta^{18}\text{O}$ values of plant water are shown in Fig. 3. We compared the $\delta^{18}\text{O}$ values of root/stem and leaf water sampled during the summer (there are no summer rains in this region) to those from the same plants sampled in November after the first rain. The results show that in the summer, root/stem water was more enriched in ^{18}O (mean stem water $\delta^{18}\text{O} = +0.8\text{‰}$) and leaf water was less enriched in ^{18}O (+13.9‰ above stem water value, on average) than in early winter (−2.8‰ and +20.0‰ for mean stem water and leaf water enrichment, respectively).

There was a clear discrepancy between the plant water and the soil water dug out from among the roots (even if a possible deuterium effect during uptake is considered¹⁴). In fact, the flood water, to which all the lines lead, seems to be the likely source of water for the plants found at the study site (Fig. 1). These

results clearly indicate that the highly saline soil water dug out along the main root systems was not used by the plants; to our knowledge, they provide the first evidence for the absence of a simple relationship between the salinity at any part of the root zone and the water actually used by the plants.

Plants may take up water at different locations and during different times, but it is clear that the various plant species used in this study were able to discriminate and use flood water for their growth and survival. These observations may be explained by spatial variations in soil water composition and the existence of flood water 'spots'^{3,22}. But they can also be explained by episodal or temporal variations in water uptake, so that only low salinity flood water is taken up when it is available. The salinity of the flood water available to the plants could not be directly determined, but low salinity near the surface, increasing with depth, would be expected because of the effects of leaching.

These results are extended by those presented in Fig. 3. The small enrichment of root/stem water in the dry summer is likely to reflect evaporation from the plant itself through, for example, lenticels in the stem or due to translocation from leaves. Such an effect would result when the transpiration flux through the plant is minimal, as indeed is indicated by the diminished stem-

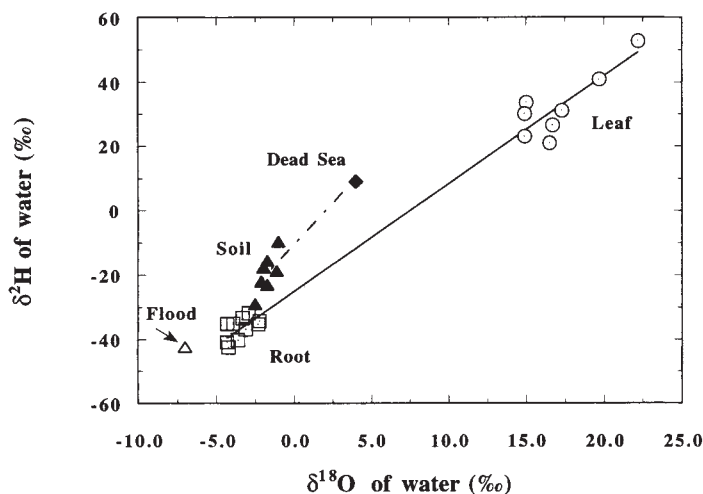


FIG. 2 Plot of the oxygen versus hydrogen isotope composition of water from leaves (○) or roots (□) of perennial plants found in the study area (Fig. 1), from soil extracted from the active root zone of these plants (▲), from the Dead Sea (◆), or of winter flood water sampled in a nearby river bed (△). Linear regression lines are plotted separately for the soil and Dead Sea water, and for the root and leaf water samples (the latter represents a group of individual lines for the different plants). Soil water was extracted by centrifugation to exclude nondissolved salts in the sediments. Water from plant tissues was cryogenically extracted under vacuum. The isotopic composition of water subsamples was determined by conventional methods of equilibrium with CO_2 (for $\delta^{18}\text{O}$) and uranium reduction (for $\delta^2\text{H}$), and by analysis of CO_2 and H_2 on an isotope ratio mass spectrometer (Finnigan MAT 250; precision was 0.2‰ and 1‰ for $\delta^{18}\text{O}$ and $\delta^2\text{H}$ determinations respectively).

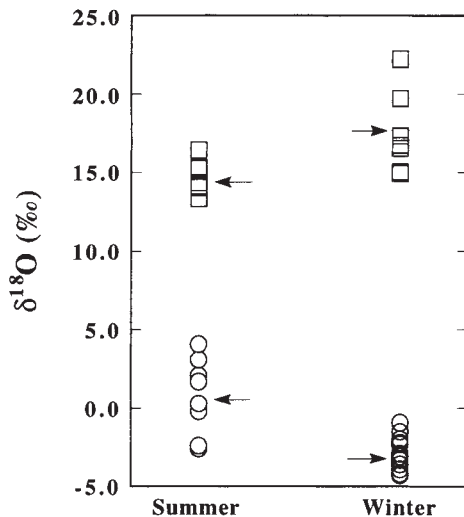


FIG. 3 Seasonal variations in the oxygen isotopic composition of water extracted from suberized root/stem (○) and leaves (□) of pioneer perennial plants found in the study area (Fig. 1). Tissue samples were collected in summer (May, no summer rains) and autumn (November, after the first rain in the region was recorded). The arrows indicate the average $\delta^{18}\text{O}$ value for leaf or root/stem water for all plant species sampled. Isotopic composition was analysed as for Fig. 2.

to-leaf enrichment—in contrast to the expected effects of the lower humidity at that time²³. Thus, the results in Fig. 3 indicate that in summer the plants become increasingly isolated from their environment, in terms of water uptake and evapotranspiration²⁴. After the first winter rain and possibly flood, the renewed utilization of fresh water results in a decrease of stem water $\delta^{18}\text{O}$, and the consequent restoration of transpiration is associated with increased isotope enrichment in the leaves. Yet it is still remarkable that the plants were capable of discriminate use of flood water, which is chemically distinct from the bulk of soil water found in their root zones. Such a feature, shared by several different plant species, may be a prerequisite for plant invasion of highly stressful environments, allowing for the early onset of terrestrial biospheric activity. □

A brain-damaged patient with an unusual perceptuomotor deficit

Umberto Castiello*†, Marina Scarpa‡
& Keree Bennett§

* Dipartimento di Psicologia, Università di Bologna,
Viale Berti Pichat 5, 40127, Bologna, Italy

‡ Clinica Neurologica, Università di Modena, Via del Pozzo, 71,
41100, Modena, Italy

§ European Medical Centre, Pieve di Cento, Bologna, Italy

† To whom correspondence should be addressed. Address from June 1995: Department of Psychology, Cippisland Campus, Monash University, Churchill, Victoria, Australia

WHEN interacting with objects, the pattern of movements is influenced by such object characteristics as size and position^{1–4}. Little is known about the effect of higher level categorical encoding of objects upon movements. Here we present evidence for an approval-for-action process which takes into account such encoding. For the brain-damaged subject L.P., the ability to complete actions involving two objects in central vision is influenced by the semantic or functional relationship between the objects. Even though she

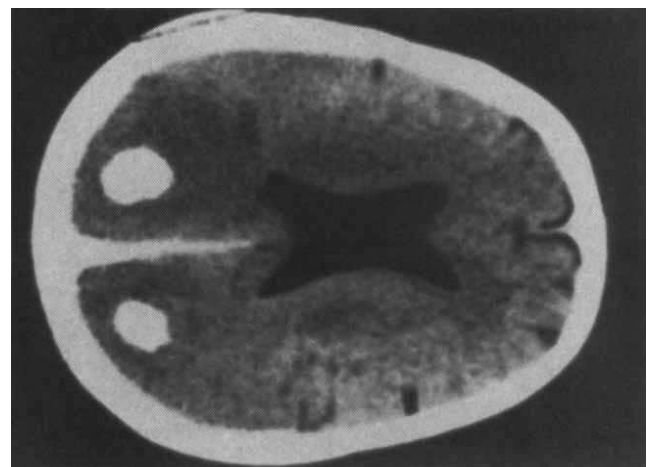


FIG. 1 L.P. suffered a left-hemisphere stroke with transient right hemiparesis in May 1993. A computed tomography (CT) scan at that time showed a left occipital lesion. The CT scan shown here demonstrates bilateral lesions (cortical mark is a slide crease). It was performed in April 1994 after L.P. complained of difficulties when putting objects together, and of intermittent problems with dressing, reading, writing and watching television. 'As my pen approaches my grandson's scribbles, the tip disappears and I cannot draw. However I can write my name'. 'At times I construct complicated LEGO forms. Yet, sometimes I cannot join pieces'. She reports: 'I see only the bottle or its lid but can screw on the lid; only the toothbrush or the tube but can squeeze out the toothpaste'. Across 15 assessment sessions during the six-month period following the second CT scan, the subjective reports and clinical picture have remained constant. L.P. identifies single elements rather than the whole of a complex picture ('telegraph boy'¹⁷; 'cookie theft'¹⁸). When presented with two pictures or objects she immediately reports seeing only one. (Note, however, that exposure time was not manipulated). There are no visual field defects (Goldman Perimeter), voluntary ocular movements are normal, and visual acuity is 10/10. L.P. is able to read single words (80/80^{18,19}) and can identify single objects: 260/260 for objects/non-objects²⁰, 99/99 for objects viewed from different angles⁷, and 32/32 for objects with similar functions but different forms²¹. L.P. has no signs of apraxia or optic ataxia^{11,22,23}. She performs common hand gestures and facial expressions, and demonstrates and mimes the use of everyday objects. She does not misreach or show inappropriate grasp apertures for a variety of targets. There are no signs of unilateral neglect, extinction, split-brain syndrome, dementia or psychiatric dysfunction.

Received 10 October 1994; accepted 20 February 1995.

- Klein, C. *Scientific Basis for Water Resources Management* (ed., M. Diskin) Vol. 2, 197–224 (IAHS publication no. 153, Oxford, UK, 1985).
- Aloni, E. & Waisel, Y. *Proc. Third Plant Life of South-west Asia Symposium* (eds Frey, W. & Kurschner, H.) (Freie Universität, Berlin, 1990).
- Shatnay, M., Magaritz, M. & Gat, J. R. *J. Coastal Res.* **4**, 257–272 (1988).
- Levitt, J. *Responses of Plants to Environmental Stress* Vol. 2, 607 (Academic, New York, 1980).
- Waisel, Y. *Biology of Halophytes* (Academic, New York, 1972).
- Flower, T. J., Hajibacheri, M. A. & Clippson, N. J. W. *Q. Rev. Biol.* **61**, 313–337 (1986).
- Zohary, M. & Orshansky, G. *Palestine J. Bot.* (Jerusalem Series) **4**, 177–206 (1949).
- White, J. W. C. in *Stable Isotopes in Ecological Research Ecological Studies* 68 (eds Rundel, P. W., Ehleringer, J. R. & Nagy, K. A.) 142–162 (Springer, Berlin, 1989).
- Sternberg, L. S. L. & Swart, P. K. *Ecology* **68**, 1898–1905 (1987).
- Sternberg, L. S. L., Ish-Shalom-Gordon, N., Ross, M. & O'Brien, J. *Oecologia* **88**, 305–310 (1991).
- Dawson, T. E. & Ehleringer, J. R. *Nature* **350**, 335–337 (1991).
- Mensforth, L. J., Thorburn, P. J., Tyerman, S. D. & Walker, G. R. *Oecologia* **100**, 21–28 (1994).
- Ehleringer, J. R. & Dawson, T. E. *Plant Cell Environ.* **15**, 1073–1082 (1992).
- Lin, G. & Sternberg, L. S. L. *Oecologia* **90**, 399–403 (1992).
- Bentor, Y. K. *Geochim. cosmochim. Acta* **25**, 239–260 (1961).
- Neev, D. & Emery, K. O. *Geol. Surv. Isr.* (bulletin no. 41, 1967).
- Gat, J. R. *Earth planet. Sci. Lett.* **71**, 361–376 (1984).
- Horita, J. & Gat, J. R. *Geochim. cosmochim. Acta* **53**, 131–133 (1989).
- Yechieli, Y. thesis, Weizmann Inst. Israel (1993).
- Allison, G. B., Gat, J. R. & Leaney, F. W. *Chem. Geol.* **58**, 145–156 (1985).
- Yakir, D., De Niro, M. J. & Gat, J. R. *Plant Cell Environ.* **13**, 49–56 (1990).
- Brendtsson, R. & Chen, H. *J. arid Environ.* **27**, 127–139 (1994).
- Yakir, D. et al. *Geochim. cosmochim. Acta* **58**, 3535–3539 (1994).
- Yakir, D., Ting, I. & DeNiro, M. J. *J. Plant Physiol.* **144**, 607–612 (1994).

ACKNOWLEDGEMENTS. We thank J. A. Raven for his comments; J. R. Gat and S. Lev-Yadun for help with sample collection and for discussion; and E. Negriano for technical assistance. D.Y. holds the Rowland Schaefer Career Development Chair.

perceives only one object, she can integrate two related objects into a coordinated action. If the objects are not related she is unable to integrate them into a single motor act. We propose that selection-for-action systems⁵ include processes which gate conceptually the behavioural disposition to action.

L.P. is a 66-year-old retired primary-school teacher with bilateral, calcarine sulcus, occipital lesions (Fig. 1). She reports seeing only one of two objects placed within central view. The selected object is correctly identified irrespective of its semantic category, or visual field position. In addition to this form of simultanagnosia^{6,8}, her ability to point to, or manipulate together, two objects is disturbed when the objects are not functionally or categorically related.

When the objects are related, her movements are performed in a unified manner, with coordinated manipulation appearing to be normal, and independent of perceptual judgements. This applies to objects that she would frequently put together. For example, when a pen and its cap are in central view she reports seeing only one or the other, yet she can pick up both objects and join them together. Unified bimanual movements are also found for pairs of three-dimensional objects, or pictures of objects, within the same 'living things'^{9,10} semantic category. If

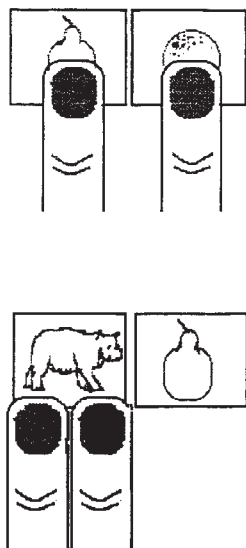


FIG. 2 Pointing task. L.P. was asked how many pictures (objects) she could see. Upon replying, she was instructed to point to them. Two pictures, pasted onto cards (3 cm × 3 cm), were presented side by side, or one above the other, with no overlap and in central view. The semantic relationship between the cards was manipulated. From 40 pictures²⁰, the following pairs were presented: 15 'living semantic'^{9,10}, from 5 cards for each animal, vegetable and fruit category; 15 'non-living' semantic^{9,10} from 5 cards for each tool, furniture and vehicle category; 30 non-semantic; and 5 associative, non-semantic (for example, grapes/wine barrel). Five pairs, presented one on top of the other, were formed for each 'living', 'non-living' and non-semantic category for objects ($n=20$) within the subject's home (for example, fruit and toys). All pairs and the following experimental design were used for all assessed tasks. Three trials were tested for each pair. Presentation order followed an ABBA design. The position of each picture (object) was randomized. Exposure time was unlimited but L.P. always responded immediately. Upper panel. 'Living' category. L.P. points to both fruits (pear and orange) yet reports seeing only one. This same result applied for 100% of trials for pictures and objects from this category. She often asked how she could point to two if she sees only one. Lower panel. No semantic relationship. L.P. points to and reports seeing only one card (bear or pear). She pointed to the bear for one trial and to the pear for two trials. Similar results were found for pairs from the 'non-living' and associative categories. The exception was when L.P. pointed to both the 'bed' and 'pillow' pictures, indicating that 'non-semantic' objects which are usually linked together in a motor act can be integrated at the motor level.

two related pictures are placed together in central view, she uses the left index finger to point to one and the right index finger to point to the other (Fig. 2, upper panel). With the index fingers thus placed, she can separate the pictures, moving both. When the pictures are placed apart she can move them together (Figs 3 and 4, upper panels).

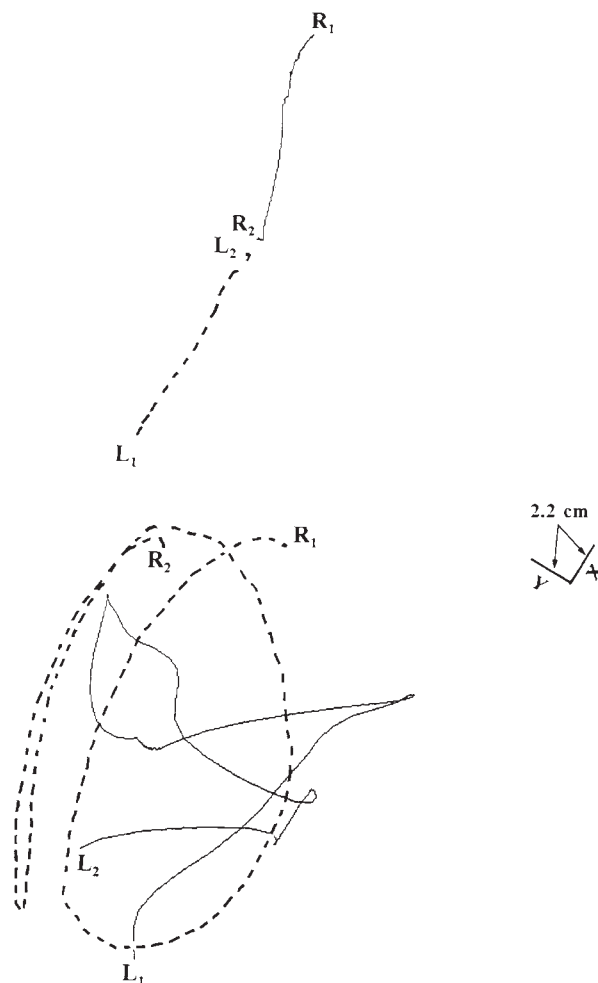


FIG. 3 L.P. was asked to put the pictures (objects) together. For bimanual movements, one was to the left and one to the right of a central point (15–60 cm apart). For unimanual movements, one was central and the other was to the right or left (10–30 cm). The subject was seated and the video view was from approximately 1 m to her left and 0.5 m above her head. The experimental design is explained in the legend of Fig. 2. For each video frame, a reference point was placed on the centre of each card (Videotrack, BTS). Broken line, left card trajectory; solid line, right card trajectory. L1 and R1 are initial positions of left and right cards; L2 and R2 are final positions. Upper panel. 'Living' semantic category. The subject places the cards (bear and alligator) together, and reports seeing one after completion of the task. L.P. also performs unimanual movements without difficulty. Lower panel. No semantic relationship. The subject begins by picking up each card (bed and lemon), but cannot bring the cards together, moving them back and forth instead in an iterative manner. Trajectories of the card movements are not consistent from trial to trial. With unimanual movements, L.P. cannot bring a laterally placed card to a central card. These results apply for all 'non-living' semantic pairs (except for bed and pillow) and for all associative non-semantic pairs. L.P. was also filmed putting together ten objects with two parts (for example, pen and pen cap), and when putting something on her finger (for example, a ring). Again, the relationship between the component parts was manipulated to form ten 'usual' pairs and ten 'unusual' pairs (for example, pen cap and tablet bottle). Each action was performed ten times. Although dextrous with 'usual' pairs, she cannot bring together 'unusual' pairs.

With object pairs that have other associative and categorical relationships her movements are completely different. She cannot bring together objects that do not usually go together. For example, she cannot complete the action of placing a pen cap on a tablet bottle; the actions taken to avoid colliding with the bottle nevertheless indicate implicit knowledge of its presence. She cannot insert her finger into a bottle top, despite correctly aligning both the finger and the top. Unified movements are also impossible with pairs of objects or pictures within the same 'non-living'^{9,10} semantic category, and with those which have no semantic relationship. They are also impossible for pairs that do not usually go together but have an associative, non-semantic relationship. When pointing to picture pairs of these three kinds, L.P. uses both index fingers to point to one or the other picture (Fig. 2, lower panel). If the examiner places L.P.'s index fingers

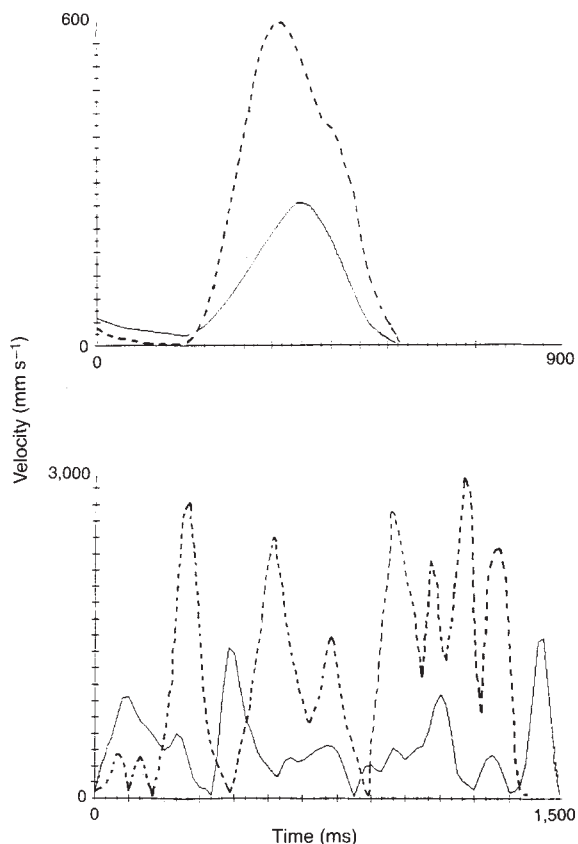


FIG. 4 Velocity profiles of the movements shown in Fig. 3. The Elite²⁴ processing package was used to filter the Videotrack data, and to construct velocity profiles. Upper panel. When bringing together pictures of the bear (left hand; broken line) and alligator (right hand; solid line), the movement of each card shows a single velocity peak. The similar timings of the velocity peaks and of the end of the card movements, suggests that the limbs have been coupled successfully for the execution of a common task²⁵⁻²⁷. Lower panel. In contrast, the bimanual action of bringing together pictures of a bed (left hand; broken line) and a lemon (right hand; solid line) demonstrates a completely different pattern. As indicated by the presence of several velocity peaks, the movement of each card is divided into several submovements. The times at which the peaks of velocity occur for the left hand do not coincide in time with those for the right hand. This, together with the asynchrony in completing the actions, suggests that the hands are acting independently, without bimanual coordination or integration. Also of interest is the high amplitude of the velocity peaks: each submovement is performed extremely rapidly. (Note that the abscissa and ordinate axes differ for the upper and lower panels). The time taken to complete the entire movement is also very long. In fact, L.P. would often continue to move the unrelated card or object pairs about in a random and disorganized manner, until taking an apparently arbitrary decision to finish the movements.

on different pictures, L.P. moves only one picture when asked to separate them. She cannot bring two separated pictures together at all (Fig. 3, lower panel): she begins the bilateral movement, but then moves the pictures in an iterative (Fig. 4, lower panel), almost circular, manner, to finish up placing the pictures often near their original positions. The uni-manual movement of bringing a laterally placed picture or object to lie beside one that is central is also not possible.

Dysfunction is present even at a representational level. To test this, picture pairs were placed face down and apart, and the examiner stated the identity of each. Despite seeing identical blank green cards, and not having seen the pictures, L.P. cannot bring together unrelated pairs.

Dysfunction is confined to central vision. While fixating centrally, L.P. correctly states the number and identity of peripherally presented objects, irrespective of their categorical or functional relationship. She can also separate them or bring them together in peripheral vision (10–30°).

This perceptuomotor deficit is characterized by visuomotor dissociations between objects in central view. A constant dissociation is for the process of consciously seeing the objects. A dissociation that is not constant, but is dependent upon the relationship between the objects, is for the performance of integrated motor acts with the objects. When the relationship is functional, motor dissociation is not present. Although this suggests preservation of familiar actions, such an explanation cannot account for the absence of motor dissociation when the objects share a 'living things' semantic category, that is, for objects (or pictures) that may not normally be put together (for example, bear and alligator). This result, together with the presence of motor dissociations for other relationships between objects, suggests that high-level visual processes (for example, semantic categorization⁷, assessment of object use and sense¹¹, and evaluation of relationships between objects) influence covert attention to the objects and motor unification of the objects.

'Implicit' or 'covert' knowledge also appears to influence attentional and motor unification. L.P. consciously uses two objects together, even though she does not report seeing both objects. (This dissociation parallels that of an agnosic patient who could act accurately upon objects using visual properties which she did not report seeing¹².) Further, and without underestimating the contribution from previous visuomotor experience, L.P. does not need low-level visual processing for unification. Even when identification of the objects is indicated verbally, conceptual gating according to the relationship between the objects continues to influence attentional and motor performance.

These results support a complex role for the occipital region in the synthesis of categorical, functional and spatial information about objects within a head-centred coordinate representation, and in the linking of this information, via inferotemporal and parietal areas^{13,14}, to motor systems^{15,16}. This selection-for-action includes a definition of the relationship between objects to be integrated within a unified motor act. □

Received 8 December 1994; accepted 20 February 1995.

1. Jeannerod, M. in *Attention and Performance* Vol. IX (eds Long, J. & Baddeley, A.) 153–168 (Erlbaum, Hillsdale, 1981).
2. Jakobson, L. S. & Goodale, M. A. *Exp Brain Res.* **86**, 199–208 (1991).
3. Castiello, U., Bennett, K. M. B. & Stelmach, G. E. *Exp Brain Res.* **94**, 165–180 (1993).
4. Weir, P. L. in *Insights into the Reach to Grasp Movement* (eds Bennett, K. M. B. & Castiello, U.) 129–150 (North-Holland, Amsterdam, 1994).
5. Allport, A. in *Perspectives on Perception and Action* (eds Heuer, H. & Sanders, A. F.) 395–419 (Erlbaum, Hillsdale, 1987).
6. Kinsbourne, M. & Warrington, E. K. *Brain* **86**, 697–702 (1963).
7. Coslett, H. B. & Saffran, E. *Brain* **114**, 1523–1545 (1991).
8. Humphreys, G. W. & Price, C. J. *Cogn. Neuropsychol.* **11**, 393–434 (1994).
9. Warrington, E. K. & Shallice, T. *Brain* **107**, 829–853 (1984).
10. Laiacina, M., Barbarotto, R. & Capitani, E. *Cortex* **29**, 727–740 (1993).
11. Jeannerod, M., Decety, J. & Michel, F. *Neuropsychologia* **32**, 369–380 (1994).
12. Milner, A. D. et al. *Brain* **114**, 405–428 (1991).
13. DeYoe, E. A. & Van Essen, D. C. *Trends Neurosci.* **11**, 219–226 (1988).
14. Goodale, M. A. & Milner, A. D. *Trends Neurosci.* **15**, 20–25 (1992).
15. Goldman-Rakic, P. S. A. *Rev. Neurosci.* **11**, 137–156 (1988).
16. Andersen, R. A. A. *Rev. Neurosci.* **12**, 377–403 (1989).

17. Kinsbourne, M. & Warrington, E. K. *Brain* **85**, 461–486 (1962).
18. Goodglass, H. & Kaplan, E. *The Assessment of Aphasia and Related Disorders* (Lea and Febiger, Philadelphia, 1972).
19. Kaplan, E., Goodglass, H. & Weintraub, S. *Boston Naming Test. Experimental Edition* (Boston Univ. Aphasia Research Center, Boston, 1982).
20. Snodgrass, J. G. & Vanderwart, M. J. *exp. Psychol., hum. Learn. Mem.* **6**, 174–215 (1980).
21. Warrington, E. K. & Taylor, A. M. *Perception* **7**, 695–705 (1978).
22. De Renzi, E. in *Neuropsychological Studies of Apraxia and Related Disorders* (ed. Roy, E. A.) 45–64 (North-Holland, Amsterdam, 1985).
23. Perenin, M. T. & Vighetto, A. *Brain* **111**, 643–674 (1988).
24. Ferrigno, G. & Pedotti, A. *IEEE Trans. biomed. Engng.* **32**, 943–950 (1985).
25. Kelso, J. A. S., Holt, K. G., Kugler, P. N. & Turvey, M. T. in *Tutorials in Motor Behavior* (eds Stelmach, G. E. & Requin, J.) 49–70 (North-Holland, Amsterdam, 1980).
26. Arbib, M. A. in *Handbook of Physiology, Section 1: The Nervous System, Vol. 2: Motor Control* (ed. Brooks, V. B.) 1449–1480 (Williams and Wilkins, Baltimore, 1981).
27. Jeannerod, M. *The Neural and Behavioural Organization of Goal-Directed Movements* (Oxford Science Publications, No. 15, Clarendon Press, Oxford, 1988).

ACKNOWLEDGEMENTS. We thank L.P. and her family, A. Allport, C. Mucignat, A. Caretta, P. Cubelli, L. Jakobson, A. Berti, M. Jeannerod, M. Goodale and A. Mazzucchi are thanked for advice and suggestions. We also thank Bioengineering Technology and Systems, Milan, Italy for the use of Elite and Videotrack.

Independent neural mechanisms for bright and dark information in binocular stereopsis

Julie M. Harris*[†] & Andrew J. Parker*

* University Laboratory of Physiology, Parks Road, Oxford OX1 3PT, UK

[†] The Smith-Kettlewell Eye Research Institute, 2232 Webster Street, San Francisco, California 94115, USA

EARLY visual processing is organized into a number of independent channels. In the retina, increments and decrements of brightness are processed independently by different groups of neurons¹. For psychophysical measurements of human vision, independence can be tested statistically. Using this criterion in a depth judgement task, we show here that, for binocular stereo vision, increments and decrements are treated independently, at least as far as the level at which information from the left and right eyes is first combined. At later stages of stereo processing, the information from the two channels is no longer independent. Because the signals for stereo vision are first combined at the visual cortex, these results suggest that the neural 'on' and 'off' channels remain independent right up to early cortical stages. Theoretical studies of stereo vision have proposed that visual features in the views of the two eyes are matched on the basis of 'similarity'². Our results show that stereo matching treats features as statistically independent (and therefore dissimilar) if they appear perceptually bright and dark relative to the background. If features differ perceptually but only in the degree of brightness or darkness, human stereo vision treats them as similar.

In the retina and the visual thalamus, the perceptual detection of increments and decrements of brightness is mediated by two independent sets of cells, the 'on' and 'off' cells^{3,4}. There is physiological evidence that the two channels may be functionally segregated as far as the cortex: however, the anatomical^{5,6} and psychophysical^{7–10} evidence is less clear. We developed a new psychophysical test of whether brightness increments and decrements are processed separately at a cortical level. To support depth judgements with binocular stereo vision, information from the two eyes is first combined effectively at the primary visual cortex. We therefore required: (1) a stimulus and task that necessitated the merging of information from the two eyes before judgements could be made and (2) a measure that reflected how much of the available information was used by the observer when the stimulus contained either just one or both contrast polarities.

The use of a random dot stereogram¹¹ satisfies the first requirement. The stereogram consists of a pair of images containing randomly positioned dots. The two eyes' images are identical, except for small horizontal offsets of the dots. With each eye viewing one half-image, the horizontal disparities between the positions of the dots are interpreted by the visual system as if the dots were at different depths. If a task requires a stereo depth judgement, it tests performance at or beyond the stage at which stereo matching takes place, as the depth information is available only after binocular combination. In the particular task studied here, the stereogram depicted a step-change in depth that appeared as a vertical edge to the observer: the observer had to identify which side of the edge was nearer (Fig. 1).

The second requirement can be satisfied by measuring the statistical efficiency of stereopsis. The measure of efficiency is essentially a comparison between the performance of a chosen system and that of an ideal observer, whose performance can be manipulated by adding noise to the stimulus¹². Efficiency gives a measure of the proportion of image information used by the test system. Thus for a dot stimulus, efficiency provides an estimate of the effective number of dots used to solve the task.

In previous work, we measured stereo efficiency using a random dot stereogram that depicted a step-edge in depth^{13,14}. Noise was added by jittering the disparity of each dot. Such noise limits the ability of even an ideal detector to decide which side of the step-edge has the greater average disparity. For this task, human efficiency is low because of an inability to exploit all the information within the target¹³. With this type of random dot stimulus that has a noisy, jagged depth profile, one main reason for the low efficiency may be that only a few dots in the pattern are correctly matched between the left and right eyes¹⁴. The problem faced by the visual system is that there are many alternative pairings of individual dots in the right and left images, most of which need to be discarded as incorrect. If the correct dots are not matched binocularly, their disparity signal is never made available to later stages of processing and thus efficiency will be low.

We considered that, if brightness increments and decrements are processed separately at the dot-matching stage, stereo efficiency should be improved by making half the dots dark and half light, because for each distinct and independent set of dots there will be fewer possible matches overall and therefore fewer incorrect matches to be discarded. With completely independent binocular matching of increments and decrements, stereo efficiency should improve by a factor of 2. If less than a factor of 2 is found, this would indicate a partial dependence: either a degree of statistical correlation between the channels or some other form of interaction.

Statistical efficiency was measured for the detection of a step-edge in depth (Fig. 1). In the first experiment, half the dots were lighter and half darker than the grey background. Performance with these mixed targets was compared against performance with stereograms containing only light or only dark dots. Figure 2a shows the effective number of dots used by an observer when the stimulus contained either dots of one contrast polarity, or dots of both polarities. Observers behaved as if they were using more dots when the stimulus contained both light and dark dots than either when only light or when only dark dots were present. Figure 3a shows the ratio of the efficiency found in the 'mixed contrast' condition to the efficiency found in the 'single contrast' condition. A ratio of 1 would indicate that the same efficiency was found in each condition, but the actual ratios are close to 2, which represents a doubling of efficiency. This suggests that the simultaneous presence of increments and decrements of brightness provided the observer with two statistically independent samples, one from the dark dots and one from the light.

In the second experiment, stimuli were created in which all dots were dark relative to the background, but the dots were at two different contrast levels ('dark' and 'darker' relative to the background). For these stimuli, the effective number of dots

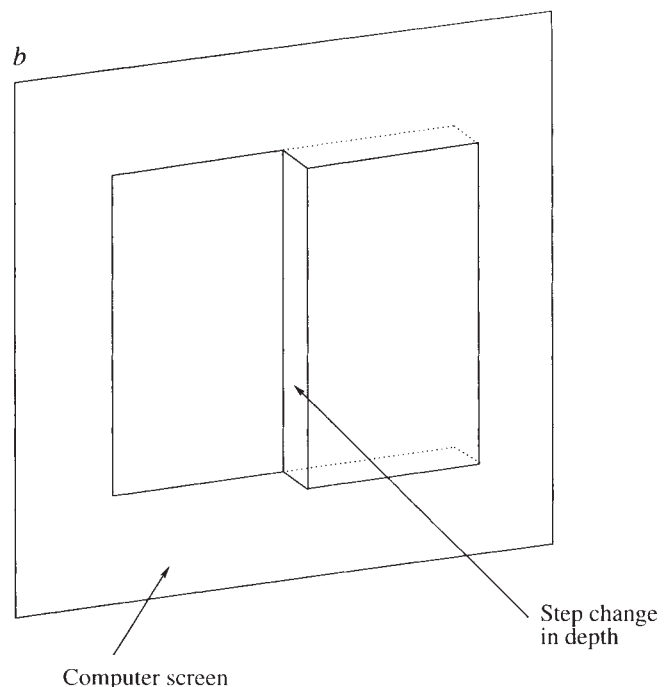
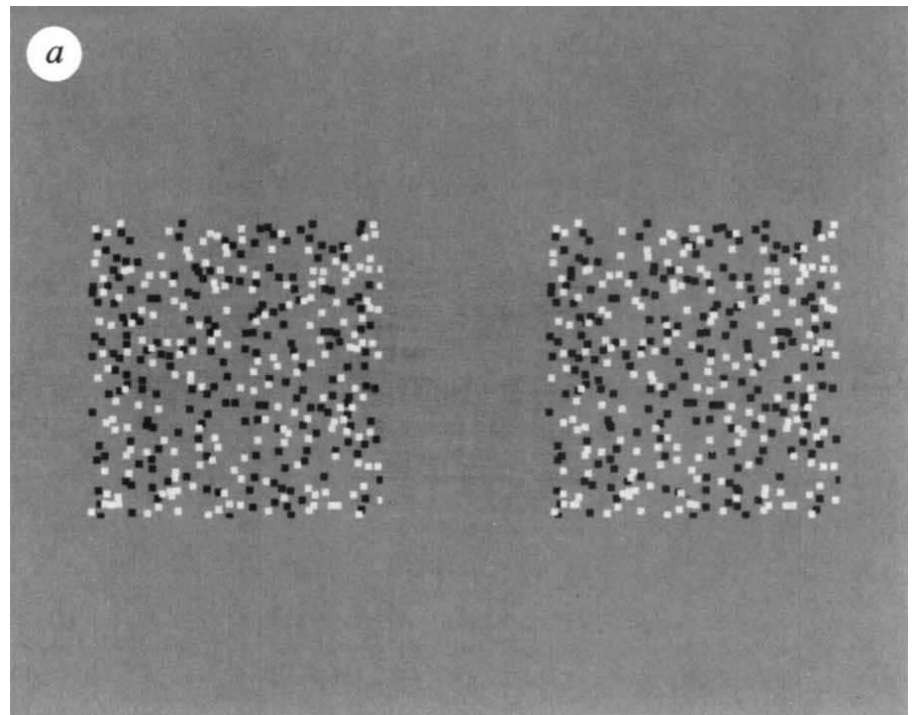
used was the same as for the patterns that contained dots of only one contrast (Fig. 2b). The ratio of efficiencies for 'mixed contrast' relative to 'single contrast' was close to 1 (Fig. 3b), suggesting that dots of the same contrast sign are not processed independently, even if they have discernibly different contrasts. Note that the difference between the results in the first and second experiments is not exclusively due to the differences in the luminance levels or magnitude of contrast between the stimuli. The peak-to-trough contrast was the same for both experiments (the contrast, $\Delta L/L_{\text{mean}}$, between the dark and the

light dots in experiment 1 was the same as the contrast between the 'darker' dots and the background in experiment 2). Additional experiments on observer J.M.H. showed that efficiency varied only slightly over a wide contrast range (Fig. 3 legend).

Are there any limitations on the statistical independence for processing increments and decrements of brightness? A third experiment indicated that later stages of processing do not exhibit complete independence. From a previous study, we know that as noise is increased, statistical efficiency gets worse¹⁴. Consider a stimulus containing dark dots with a disruptive level

FIG. 1 The random dot stereogram stimuli consisted of a square area subtending 90° (visual angle at the eye) at the 2.9 m viewing distance and contained between 60 and 600 dots (each dot of side 3.2'). The stereogram (a) portrayed a flat front parallel surface bisected by a vertically oriented step change in depth at the centre of the stimulus (a schematic version is shown as b). Noise was added by giving each dot an extra disparity drawn randomly from a normal distribution of known variance. The resulting stimulus contained dots at many different depths, but the dots on one side of the vertical edge were on average nearer in depth to the observer. The task of the observer was to identify which side of the step-edge was nearer. The disparity of the step-edge was set to be 1.4' and each sample of added noise was drawn from a distribution of standard deviation 3'. The outer left and right boundaries of the step-edge stimulus were sloped back to zero disparity¹⁴ so that there was no monocularly available cue that could unintentionally reveal the nature of the step in depth.

METHODS. There were three types of stimuli used: (1) 50% of the dots were 'dark' and 50% 'light', or (2) 100% of the dots were 'dark', or (3) 100% of the dots were 'light'. a, An example of a stimulus from condition 1, where there were 120 light dots and 120 dark dots in the stimulus. If the stereo-halves are crossfused, the stereogram will portray a cloud of dots in depth, with the right-hand side standing out further on average than the left (as depicted in b). In the first experiment, the luminance of the grey background, L_{mean} , was 56 cd m⁻². The dark dots had luminance 38 cd m⁻² and the light 71 cd m⁻², making the contrast, $\Delta L/L_{\text{mean}}$, of each against the background about 30%. In the second experiment, two different contrasts were used, but both had the same sign (both were darker than the background with luminances of 38 cd m⁻² and 20 cd m⁻²). It was noticed initially that these dots were readily discernible from one another, but were not equally salient perceptually. To improve the salience of the lower contrast dots, their size was increased slightly relative to the higher contrast dots. The higher contrast (darker) elements were set to be of side 2.5' and the lower contrast (dark) elements were of side 3.6'. With this arrangement, the two types of dot appeared equally discernible relative to the background. A single experimental run consisted of the presentation of 140 stimuli and on each trial the observer had to decide which side of the step edge stood out further in depth. No fixed time was set for viewing or response, but experienced observers generally took between 1 and 2 s for each decision. The percentage of correct responses was recorded from at least 3 runs per condition and transformed to d' using the inverse of the cumulative-normal equation¹⁴. Because we have shown previously that d' is proportional to the size of the depth step for similar experiments¹³, we were satisfied that one value of d' was representative of the d' -disparity relation. Efficiency was calculated by comparing the measured discriminability, d'_e , with that of an ideal observer, d'_i . For the ideal observer, d'_i is defined as $d'_i = \Delta M/\sigma$, where ΔM is the disparity of the depth step and σ is the standard error of the mean disparity¹³. Efficiency is defined as: $F = (d'_e/d'_i)^2$ and is a measure of the amount of information actually used compared with that potentially available. The correction of d'_i and d'_e for the sloped edges of the stimulus was done according to the procedures developed previously¹⁴.



of noise and light dots with no added noise. If observers had unhindered access to an 'on' channel transmitting only brightness increments, they should use only the light dots and be unaffected by the noise in the dark dots. This does not happen. Stereo vision is significantly disrupted when the dark dots contain a high level of noise. We systematically varied the relative levels of noise of the dark and light dots. Performance for stimuli containing both populations of dots was always intermediate between the performance found for the dark-only and light-only stimuli, suggesting that the 'on' and 'off' channels are not separate at the level where individual depth estimates are combined. Thus, although contrast polarity may be used to assist binocular matching in a population of disparity-selective neurons, the signal delivered by the output of such neurons may

reflect only the disparity values and may fail to indicate which feature generated any particular disparity value.

Human observers make use of between 1% and 27% of the dots available in the display (Fig. 2). Particularly at the lower values, this is a striking limitation on performance compared with efficiencies for other tasks^{12,15,16}. Low efficiency would be expected if, because of particular matching strategies, only a few dots were matched between the left and right eyes' images or if the disparity signal from some matches was discarded^{17,19}. Whatever the exact method by which correct matching occurs, information about the light and dark dots appears to pass through separate channels until after the stage of stereo matching. Because it is thought that binocular matching occurs within the cortex, this would imply a functional segregation of 'on' and 'off' channels until at least this point in the visual pathways. Not

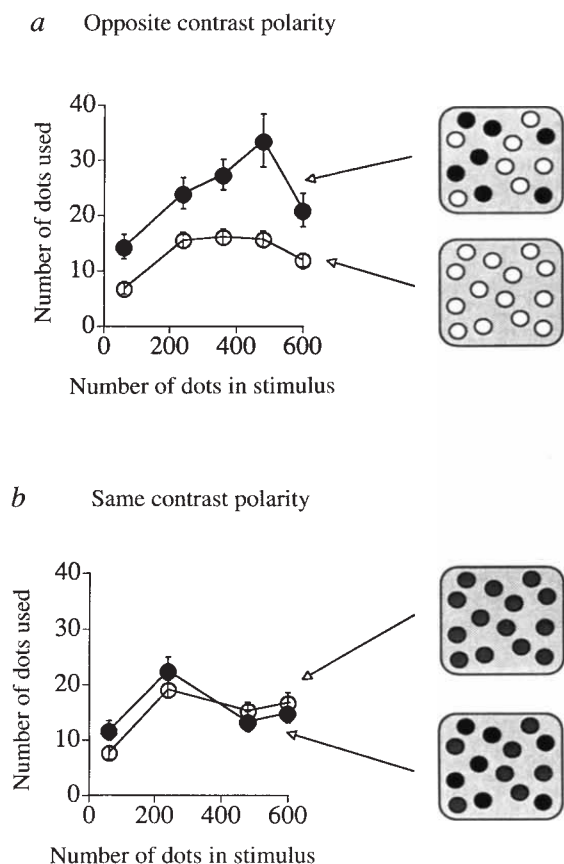


FIG. 2 The statistical efficiency of stereo depth detection was measured for stimuli containing dark and light elements (the mixed condition), and for stimuli with all dark or all light elements (single contrast conditions). Preliminary experiments showed that efficiency for the 'all dark' and 'all light' stimulus conditions was very similar. The results of the two conditions were therefore averaged in each experiment. Efficiency can be expressed in terms of the number of dots used to solve the task: effective number of dots used by observer = efficiency $\times$ number of dots in stimulus. **a**, Graph shows the effective number of dots as a function of the actual number of dots in the stimulus for the first experiment, where 'light' and 'dark' dots had opposite contrast sign. The solid symbols show the results for the stimuli containing both light and dark dots, the open symbols show the results for stimuli containing either only dark or only light dots. It is as if the observer uses almost twice as many dots when the stimuli contain both dark and light dots than when the stimuli contain dots of a single brightness. **b**, Graph shows the effective number of dots used when both types of element were darker than the background. The solid symbols show the results for the stimuli containing both dark and darker dots, the open symbols the results for stimuli containing either only dark or only darker dots. The effective number of dots used by the observers was similar, regardless of whether the stimuli contained dots of a single brightness or two different brightnesses.

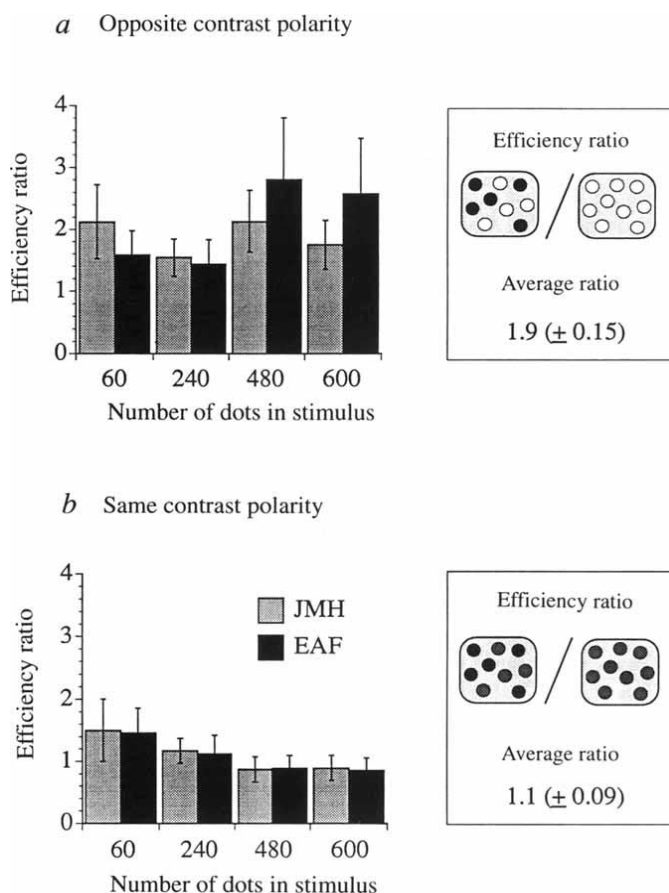


FIG. 3 The efficiency ratio is plotted for two observers for the opposite contrast polarity experiment (**a**) and the same contrast polarity experiment (**b**). The efficiency ratio is the ratio of the efficiency measured in the mixed contrast condition to the efficiency measured in the single contrast condition. When the mixed condition contained dots of opposite contrast polarity, the ratio was close to 2 for both observers, over the whole range of stimuli (from 60 to 600 dots). This suggests that twice as much information was used in the mixed polarity condition, as compared with the single polarity condition, and hence that there was independent processing of the dark and light dots. When the mixed condition contained dots of the same polarity, but different brightness, the ratio was close to 1. This suggests that the dark and darker dots were not being matched independently and were not treated in any way separately. Independent processing occurs only when the dots have opposite contrast polarity. Additional experiments established that the efficiency did not vary over a wide range of dot luminance levels (or equivalent dot contrasts relative to the background). For example, in the condition where the stimulus contained 'dark' and 'darker' dots, an efficiency ratio of close to 1 (no independence) was found for observer J.M.H. for brightness differences ranging from 18 to 35 cd m^{-2} (a contrast range of 0.2–0.8).

only do the results provide additional evidence that the stereo matching process does not match features of opposite contrast^{11,20,21}, they also demonstrate a new aspect of stereo correspondence: the elimination of false matches proceeds independently on bright and dark elements. Thus, if there are computational constraints involved in assigning binocular matches^{17,19}, the constraints cannot be purely geometrical.

These results complement work on the independence of 'on' and 'off' signals in other areas of early visual processing. One model for early encoding of spatial features relies on the separate processing of 'on' and 'off' signals and is supported by a wide range of data²². Separate processing of opposite contrast-polar-

ity signals has also been suggested for spatial interval discrimination^{23,24} and for motion processing²⁵. Again, in a similar way to our stereo data, studies of both motion²⁵ and position processing²⁶ suggest that at higher levels the processing of 'on' and 'off' signals is no longer independent.

A general conclusion from the strategy developed here is that measures of statistical efficiency can be adopted to discover the number and types of visual channel that are used by human observers. Many approaches used to identify and characterize channels have relied upon analysing vision at threshold. This work shows how the notion of independent channels can be applied to higher levels of perception that involve more natural aspects of visual processing. □

Received 25 October 1994; accepted 9 March 1995.

- Schiller, P. H. *Trends Neurosci.* **15**, 86–92 (1992).
- Marr, D. & Poggio, T. *Proc. R. Soc. B* **204**, 301–328 (1979).
- Schiller, P. H., Sandell, J. L. & Maunsell, J. H. R. *Nature* **322**, 824–825 (1986).
- Dolan, R. P. & Schiller, P. H. *Visual Neurosci.* **11**, 23–32 (1994).
- McConnell, S. K. & LeVay, S. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1590–1593 (1984).
- Zahs, K. R. & Stryker, M. P. *J. Neurophysiol.* **59**, 1410–1429 (1988).
- DeValois, K. K. *Vision Res.* **17**, 209–215 (1977).
- Burton, G. J., Naghshineh, S. & Ruddock, K. H. *Biol. Cyber.* **27**, 189–197 (1977).
- Bjorklund, R. A. & Magnussen, S. *Perception* **10**, 511–518 (1981).
- Hanly, P. & MacKay, D. M. *Expl Brain Res.* **35**, 37–46 (1979).
- Julesz, B. *Foundations of Cyclopean Perception* (Univ. Chicago Press, Chicago, 1971).
- Barlow, H. B. *Vision Res.* **18**, 637–650 (1978).
- Harris, J. M. & Parker, A. J. *J. opt. Soc. Am. A* **1**, 14–24 (1992).
- Harris, J. M. & Parker, A. J. *Vision Res.* **34**, 2761–2772 (1994).

- Burgess, A. & Barlow, H. B. *Vision Res.* **23**, 811–820 (1983).
- Barlow, H. B. & Reeves, B. C. *Vision Res.* **19**, 783–793 (1979).
- Poliard, S. B., Mayhew, J. E. W. & Frisby, J. P. *Perception* **14**, 449–470 (1985).
- Marr, D. & Poggio, T. *Science* **194**, 283–287 (1976).
- Baker, H. H. & Binford, T. O. *Proc. 7th Int. Joint Conf. Artificial Intelligence* 631–636 (1981).
- von Helmholtz, H. *Handbook of Physiological Optics* 3rd edn (1910, translated 1925 by Southall, J.) (Dover Publications, New York, 1953).
- Kaufman, L. & Pitblado, C. B. *Percept. Psychophys.* **6**, 10–12 (1969).
- Watt, R. J. & Morgan, M. J. *Vision Res.* **25**, 1661–1674 (1985).
- Levi, D. M. & Westheimer, G. *J. opt. Soc. Am. A* **4**, 1304–1313 (1987).
- Levi, D. M., Jiang, B.-C. & Klein, S. A. *Vision Res.* **30**, 1735–1750 (1990).
- Edwards, M. & Badcock, D. R. *Vision Res.* **34**, 2849–2858 (1994).
- Morgan, M. J. & Giennester, A. *Vision Res.* **31**, 2075–2083 (1991).

ACKNOWLEDGEMENTS. We thank E. Fraser for acting as an observer in these experiments. This work was funded by the SERC, the Wellcome Trust and the McDonnell-Pew Centre for Cognitive Neuroscience. J.M.H. was partially supported by a Wellcome Trust Prize Fellowship.

Prevention of hypoxia-induced cell death by Bcl-2 and Bcl-xL

Shigeomi Shimizu^{*†}, Yutaka Eguchi[‡],
Hiroaki Kosaka[§], Wataru Kamiike^{*},
Hikaru Matsuda^{*} & Yoshihide Tsujimoto[‡]

^{*} Department of Surgery, [†] Department of Medical Genetics, Biomedical Research Center, and [§] Department of Physiology, Osaka University Medical School, 2-2 Yamadaoka, Suita, Osaka, 565, Japan

THE proto-oncogene *bcl-2*, isolated from the t(14;18) chromosomal breakpoint in follicular B-lymphoma^{1–3}, and a *bcl-2*-related gene *bcl-x* (ref. 4) prevent apoptotic cell death induced by various treatments^{5–8}. Although a mechanism has been proposed that involves Bcl-2 activity on reactive oxygen species (ROS)^{9,10}, expression of Bcl-2 or Bcl-xL prevents cell death induced by withdrawal of oxygen (hypoxia), which drastically decreases the net formation of oxygen free radicals and does not increase oxidized lipid, protein or DNA. Furthermore, neither ROS scavenger nor inhibitor of ROS scavenger affects cell death, regardless of the expression of Bcl-2 or Bcl-xL. Thus our data suggest that Bcl-2 and Bcl-xL exert an anti-cell death function by a mechanism other than regulation of ROS activity.

To clarify whether or not the anti-cell death activity of Bcl-2 is due to its regulation of ROS, we analysed the effects of Bcl-2 on the cell death induced by depletion of oxygen (less than 100 p.p.m. O₂), where ROS involvement is unlikely, using rat hepatoma cell line 7316A (ref. 11) and rat pheochromocytoma cell line PC12 (ref. 12). The vector transfectants of these cells died gradually and, after 48 hours, 50% to 60% had survived (Fig. 1a, b, open symbols), as measured by trypan blue exclusion. Cell death observed here surely resulted from hypoxia, and not reoxygenation during cell survival analysis, because death by reoxygenation could not be observed within 1 hour (not shown).

Overexpression of human Bcl-2 (Fig. 1e, f) significantly increased survival of these cells (Fig. 1a, b, filled symbols), indicating that Bcl-2 prevented cell death induced by withdrawal of oxygen in 7316A and PC12 cells.

Bcl-xL is similar in size and predicted structure to Bcl-2, and prevents apoptotic cell death⁴. Overexpression of Bcl-xL protein in PC12 cells (Fig. 1g) prevented cell death induced by hypoxia (Fig. 1c). Similar survival curves were obtained when the extent of cell death was measured by the release of lactate dehydrogenase, a cytoplasmic enzyme (Fig. 1d). Bcl-xL was more efficient than Bcl-2 in preventing cell death (cf. Fig. 1b, c), but it is not clear whether this reflects the difference of their activities or amounts.

The efficiency of the formation of ROS would be greatly reduced under hypoxic conditions, where oxygen concentrations are extremely low. Indeed, analysis of electron spin resonance (ESR)¹³ using PC12 vector transfectants revealed signals of the spin-trapped oxygen free radicals under normoxic conditions (Fig. 2a), but not under hypoxic conditions (Fig. 2b). A drastic reduction in the signals by the addition of Cu/Zn superoxide dismutase (Fig. 2c), but not by the addition of catalase (Fig. 2d), suggests that most of the signals derived from superoxides. Note that even the high levels of ROS under normoxic conditions are not harmful to cells. Together, our results indicate that the hypoxia-induced cell death occurs under conditions in which the ROS are hardly produced. No increase in oxidized DNA, protein or lipid was detected (Fig. 3), indeed there was a slight decrease. Therefore prevention of cell death by Bcl-2 and Bcl-xL appears to occur in the absence of ROS.

To exclude the theoretical possibility that undetectable amounts of ROS are involved in hypoxia-induced cell death, we tested whether ROS scavengers inhibit cell death during hypoxia and, conversely, whether ROS scavenger inhibitors would enhance cell death. Addition of a ROS scavenger, N-acetylcysteine, which rapidly reacts with and inactivates ROS¹⁴, significantly inhibited cell death induced by reoxygenation after the hypoxic treatment (Fig. 4a), whereas it had essentially no effect on cell death under hypoxic conditions (Fig. 4b). A ROS scavenger inhibitor, diethyl maleate (DEM), binds directly to glutathione, a tripeptide necessary for the decomposition of hydrogen peroxide by glutathione peroxidase¹⁵. Treatment of PC12 with DEM under normoxic conditions led to cell death within 48

[†] Present address: 1st Department of Physiology, Osaka University Medical School, 2-2 Yamadaoka, Suita, Osaka, 565, Japan

|| To whom correspondence should be addressed.

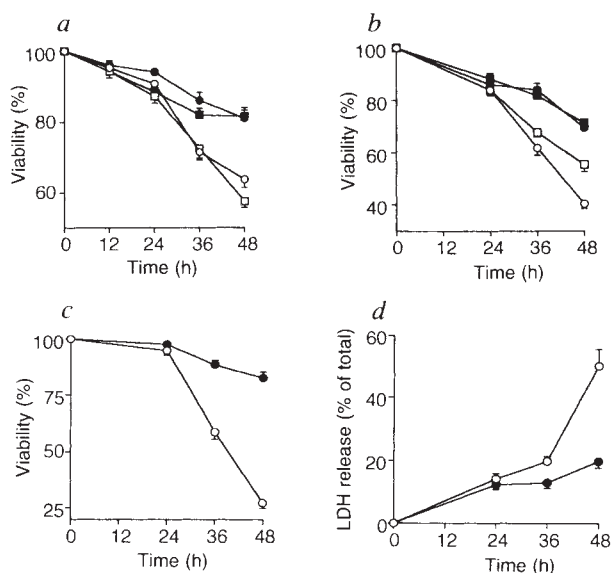
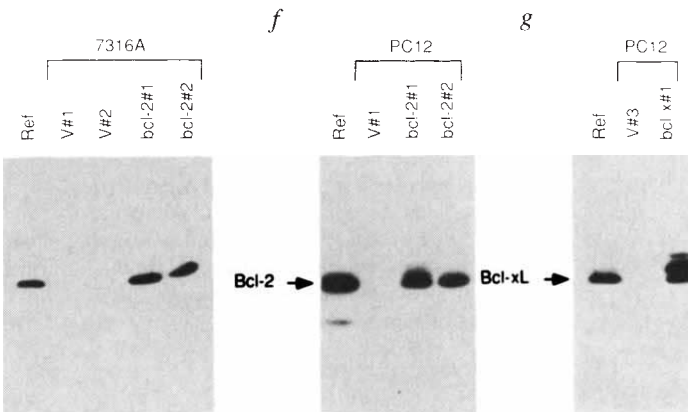
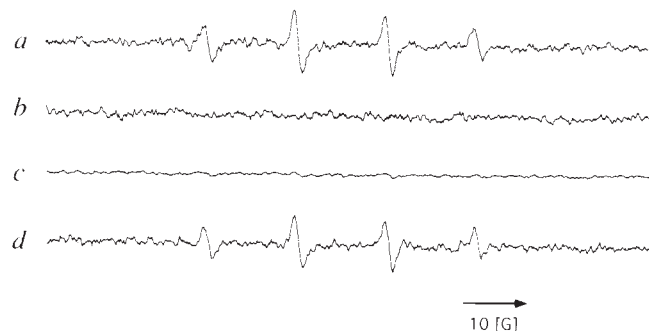


FIG. 1 Cell death by hypoxia and its prevention by Bcl-2 or Bcl-xL. **a**, Viability of 7316A derivatives expressing human Bcl-2; 7316Abcl-2 #1 (filled circles), 7316Abcl-2 #2 (filled squares), and their control clones, 7316AV #1 (open circles) and 7316AV #2 (open squares). **b**, Viability of PC12 derivatives expressing human Bcl-2; PC12bcl-2 #1 (filled circles), PC12bcl-2 #2 (filled squares), and their control clones, PC12V #1 (open circles) and PC12V #2 (open squares). **c** and **d**, Viability of PC12 derivative expressing chicken Bcl-xL (PC12bcl-x #1; filled circles) and its control clone (PC12V #3; open circles). Viability of each clone was measured by two assays, trypan blue exclusion (**a**–**c**), and lactate dehydrogenase (LDH) release (**d**), and data represent the means of 5 independent experiments $\pm$ s.d. for each clone. **e**–**g**, Expression of Bcl-2 and Bcl-xL in representative clones analysed by western blotting using *in vitro* translated proteins as references (ref). More than two bands

hours, with an LD₅₀ (the dose lethal to 50% of cells tested) of about 0.3 mM (Fig. 4c), but death was partially prevented by expression of Bcl-2 or Bcl-xL. However, DEM did not enhance hypoxia-induced cell death even at 1 mM (Fig. 4d), regardless of the expression of Bcl-2 or Bcl-xL. These observations indicate that even very small amounts of ROS are not involved in hypoxia-induced cell death.

From the data presented here, we conclude that the prevention of the hypoxia-induced cell death by Bcl-2 or Bcl-xL is not due to their regulation of ROS. Most, if not all, cell death appears to be apoptosis, as indicated by the nuclear fragmentation seen in Hoechst33342-stained hypoxic cells (S.S. *et al.*, in preparation). Another type of cell death, Fas-mediated apoptosis^{16,17}, and Bcl-2 prevents the apoptosis¹⁸. Even in ROS-mediated apoptosis,



were observed for Bcl-xL, probably resulting from some post-translational modification.

METHODS. Human Bcl-2 cDNA and chicken Bcl-xL cDNA cloned from the respective libraries were inserted in retroviral vector pBC140 (ref. 24) and were used to infect cells via the packaging cell line Ψ 2. For experiments with Bcl-xL, more than 10 transfectants were pooled. Flow cytometric analysis showed that more than 98% of cells expressed the exogenous gene products (not shown). Control clones were prepared similarly using the vector without the inserts. Cells (10^5) plated on 6-cm dishes were transferred to a chamber, and hypoxic conditions were achieved using BBL GasPac Plus (Becton Dickinson, USA), which catalytically reduces O₂ levels to less than 100 p.p.m. within 90 min (time 0 of the hypoxic conditions). At the indicated times, cells were collected, and surviving cells were counted after exposure to trypan blue, or fractions of dead cells were estimated from the ratio of the activity of LDH released in the medium to its total activity. Western blotting was performed using mouse anti-human Bcl-2 monoclonal antibody²⁵, and rabbit anti-chicken Bcl-xL polyclonal antibody that was raised against glutathione S-transferase (GST)–Bcl-xL fusion protein and then affinity purified.

we speculate that Bcl-2 or Bcl-xL functions at the same step as in the hypoxia-induced cell death and Fas-mediated apoptosis. The reported Bcl-2-induced decrease in the production of ROS¹⁰ and the Bcl-2-induced decrease in the oxidative activity of ROS without affecting its production⁹ might reflect the secondary effects of prevention of cell death by Bcl-2.

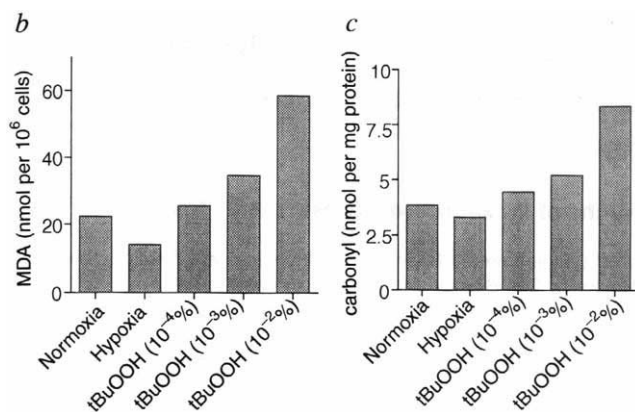
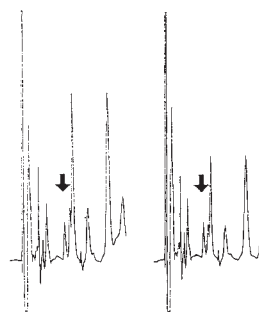
What are the targets of Bcl-2 and Bcl-xL? Increased intracellular Ca²⁺ concentrations and protease function might be critical in hypoxia-induced cell death¹⁹. This idea is consistent with the fact that the *ced-3* gene product, which is similar to a mammalian cysteine protease ICE (ref. 20), is essential for programmed cell death in *Caenorhabditis elegans*²¹. The *ced-9* gene, a homologue of *bcl-2* (ref. 22), appears to regulate the function of *ced-3* in providing protection from cell death²³. Thus Bcl-2 and Bcl-xL might prevent cell death by regulating proteases directly, or

FIG. 2 Reduction of formation of oxygen free radicals in hypoxia. Data represent the means of 3 scans of ESR spectra of spin-trapped oxygen free radicals formed in PC12V #1 cells **a**, during 1-hour incubation under normoxic conditions; **b**, during the first 1-hour incubation under hypoxic conditions; **c**, under the same conditions as **a** but in the presence of 50 units ml⁻¹ Cu/Zn-dependent superoxide dismutase; and **d**, under the same conditions as **a** but in the presence of 500 units ml⁻¹ catalase.

METHODS. Cells were mixed with spin-trapping reagents, 5,5-dimethyl-1-pyrroline N-oxide (DMPO) and transferred to hypoxic conditions. ESR detection of spin adducts was performed with a Bruker ESP300 ESR spectrometer. Culture supernatant was collected and transferred to a quartz ESR flat cell, which was in turn placed in the cavity of the ESR spectrometer. ESR settings were as follows: incident microwave power, 20 mW; modulation frequency, 100 kHz; modulation amplitude, 1.0 G; response time, 0.7 s; scan width, 100 G; and scan time, 2.8 min.

FIG. 3 No increase in oxidized DNA, lipid or protein in hypoxia. **a**, The 8-OH-2'-deoxyguanosine (8-OH-dG), a product of oxidized DNA, was quantified as described²⁶ in samples prepared from PC12 cells (5×10^7) incubated under hypoxic conditions for 48 hours (left) or under normoxic conditions (right); 0.63 and 0.79 residues of 8-OH-dG (shown by arrows) per 10^5 dG residues in hypoxia and normoxia, respectively, were detected. **b** and **c**, Malondialdehyde (MDA), a product of peroxidized lipid (**b**), and carbonyl content of proteins, an index of oxidative modification of proteins (**c**), were measured as described^{27,28} with some modifications, in samples of PC12 cells grown under normoxic conditions, hypoxic conditions for 4 hours, or, as a positive control, treated with a ROS producer, *t*-butyl hydroperoxide (tBuOOH) at indicated concentrations.

METHODS. In **b**, cells (10^6) were homogenized in 1 ml PBS. After adding 2 ml of 0.375% thiobarbituric acid/0.25 M HCl/0.06% 2,6-di-*t*-butyl-4-hydroxytoluene/22% ethanol, the mixture was incubated at 100 °C for 15 min, centrifuged at 4 °C, and the OD_{535nm} of the supernatant was measured. In **c**, cells (10^6) were homogenized in 0.5 ml lysing buffer



and passed through 0.45- μ m filters. Proteins were precipitated by trichloroacetic acid (TCA), resuspended in 0.5 ml 10 mM, 2,4-dinitrophenylhydrazine/2 M HCl and incubated for 1 hour at room temperature. Proteins were precipitated by TCA, washed with ethanol and redissolved in 0.6 ml 6 M guanidine/10 mM phosphate buffer/trifluoroacetic acid, pH 2.3. After 15-min incubation at 37 °C, the OD_{366nm} of the supernatant was measured. Carbonyl content was calculated using an extinction coefficient of 22.0 mM⁻¹ cm⁻¹ for aliphatic hydrazones.

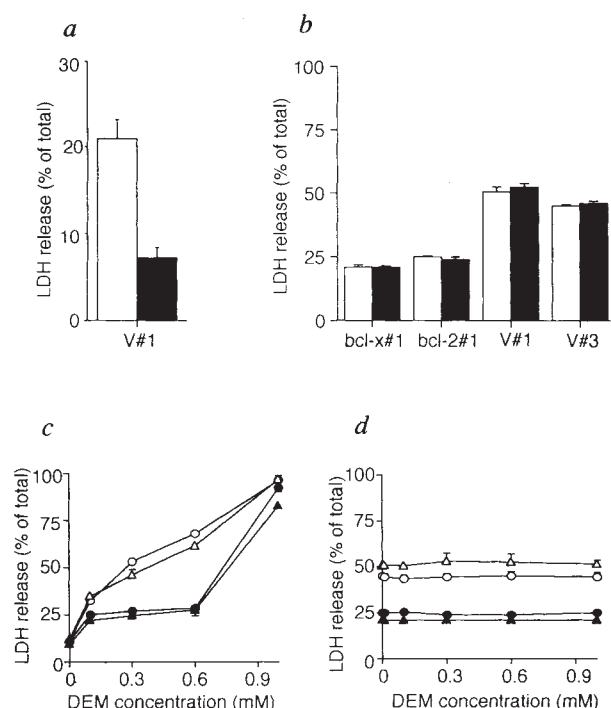


FIG. 4 Cell death of PC12 in response to ROS scavenger and inhibitor of ROS scavengers in the presence or absence of oxygen. **a**, Effect of *N*-acetylcysteine (NAC) on the death of PC12V#1 induced by reoxygenation after 24 hours of hypoxia treatment. **b**, Effect of NAC on the death of PC12bcl-2#1, PC12bcl-x#1, PC12V#1 and PC12V#3 under hypoxic conditions. **c** and **d**, Cell death in response to diethyl maleate (DEM) in the presence (**c**) or absence (**d**) of oxygen: PC12bcl-2#1 (filled circles), PC12bcl-x#1 (filled triangles), PC12V#1 (open circles) and PC12V#3 (open triangles). Data represent the means of 3 independent experiments $\pm$ s.d. for each clone.

METHODS. For **a** and **b**, cells were incubated in the presence (filled bars) or absence (open bars) of 1 mM NAC under normoxic conditions for 4 hours after the 24 hours of hypoxic treatment or under hypoxic conditions for 48 hours, and the fractions of dead cells were estimated from the activity of LDH released in the medium. For **c** and **d**, cells (10^5) were plated on 6-cm dishes and DEM was added at the concentrations indicated, followed by immediate transfer to the hypoxic chamber as described in the Fig. 1 legend. After 48 hours of incubation, fractions of dead cells were estimated from the activity of LDH released in the medium.

might function through protein kinase cascades, ion channels, cytoskeletons or other small compounds. The targets of Bcl-2 and Bcl-xL are yet to be elucidated. □

Received 5 December 1994; accepted 21 March 1995.

1. Tsujimoto, Y., Cossman, J., Jaffe, E. & Croce, C. M. *Science* **228**, 1440–1443 (1985).
2. Bakhshi, A. et al. *Cell* **41**, 899–906 (1985).
3. Cleary, M. L. & Sklar, J. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7439–7443 (1985).
4. Boise, L. H. et al. *Cell* **74**, 597–608 (1993).
5. Vaux, D. L., Cory, S. & Adams, J. M. *Nature* **335**, 440–442 (1988).
6. Tsujimoto, Y. *Proc. natn. Acad. Sci. U.S.A.* **86**, 1958–1962 (1989).
7. Nunez, G. et al. *J. Immunol.* **144**, 3602–3610 (1990).
8. Vaux, D. L. *Proc. natn. Acad. Sci. U.S.A.* **90**, 786–789 (1993).
9. Hockenbery, D. M., Oltvai, Z. N., Yin, X.-M., Millman, C. L. & Korsmeyer, S. J. *Cell* **75**, 241–251 (1993).
10. Kane, D. J. et al. *Science* **262**, 1274–1277 (1993).
11. Hirota, K., Hirota, T., Sanno, Y. & Tanaka, T. *Cancer Res.* **47**, 3742–3746 (1987).
12. Greene, L. A. & Tischler, A. S. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2424–2428 (1976).
13. Takeuchi, T., Morimoto, K., Kosaka, H. & Shiga, T. *Biochem. biophys. Res. Commun.* **194**, 57–64 (1993).

14. Staal, F. J. T., Roederer, M., Herzenberg, L. A. & Herzenberg, L. A. *Proc. natn. Acad. Sci. U.S.A.* **87**, 9943–9947 (1990).
15. Ku, R. H. & Billings, R. E. *Archs Biochem. Biophys.* **247**, 183–189 (1986).
16. Schulze-Osthoff, K., Krammer, P. H. & Droge, W. *EMBO J.* **13**, 4587–4596 (1994).
17. Hug, H., Enari, M. & Nagata, S. *FEBS Lett.* **351**, 311–313 (1994).
18. Itoh, N., Tsujimoto, Y. & Nagata, S. *J. Immunol.* **151**, 621–627 (1993).
19. Khalbuss, W. E. & Wondereg, R. *Hepatology* **13**, 962–969 (1991).
20. Yuan, J., Shaham, S., Ledoux, S., Ellis, H. M. & Horvitz, H. R. *Cell* **75**, 641–652 (1993).
21. Ellis, H. M. & Horvitz, H. R. *Cell* **44**, 817–829 (1986).
22. Hengartner, M. O. & Horvitz, H. R. *Cell* **76**, 665–676 (1994).
23. Hengartner, M. O., Ellis, R. E. & Horvitz, H. R. *Nature* **356**, 494–499 (1992).
24. Borzillo, G. V., Endo, K. & Tsujimoto, Y. *Oncogene* **7**, 869–876 (1992).
25. Pezzella, F., et al. *Am. J. path.* **137**, 225–232 (1990).
26. Shigenaga, M. K., Park, J.-W., Cundy, K. C., Gimeno, C. J. & Ames, B. N. *Meth. Enzym.* **186**, 521–530 (1990).
27. Esterbauer, H. & Cheeseman, K. M. *Meth. Enzym.* **186**, 407–428 (1990).
28. Levine, R. L. et al. *Meth. Enzym.* **186**, 464–478 (1990).

ACKNOWLEDGEMENTS. We thank M. Raff and colleagues for communicating their results before publication; D. Mason for providing anti-human Bcl-2 monoclonal antibody clone 100; K. Tamai and T. Hachiya for their cooperation in generating antibodies; and M. Hoffman for editorial assistance. This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan.

Programmed cell death and Bcl-2 protection in very low oxygen

Michael D. Jacobson & Martin C. Raff

MRC Developmental Neurobiology Programme, MRC Laboratory for Molecular Cell Biology and the Biology Department, University College London, London WC1E 6BT, UK

PROGRAMMED cell death (PCD) is a fundamental feature of animal cells¹, but the mechanism remains unknown. Similarly, the Bcl-2 oncoprotein can suppress PCD in a variety of cell types and circumstances², but it is not known how it does so. It has been suggested that PCD involves the generation of reactive oxygen species (ROS) and that Bcl-2 protects against PCD by inhibiting the generation or action of ROS³⁻⁶. To determine whether ROS are required for PCD, we cultured cells in a near-anaerobic atmosphere where the generation of ROS would be expected not to occur, or at least to be greatly reduced. We find that these conditions inhibit PCD induced by ROS-generating agents but do not inhibit PCD induced by other means. Furthermore, we show that Bcl-2 can protect cells from PCD in these anaerobic conditions. These results suggest that ROS are not required for PCD, and that Bcl-2 protects against PCD in ways that do not depend on the inhibition of ROS production or activity.

Cells were cultured in a mixture of 10% CO₂ and 90% N₂ in a humidified 37 °C incubator contained within a sealed, anaerobic, gloved cabinet fitted with two heated copper catalyst columns to scavenge free oxygen. Oxygen levels in the cabinet were measured with a monitor that was sensitive to oxygen concentrations down to <1 p.p.m. and was accurate to within $\pm 3\%$. Cells were cultured in these conditions for 24 hours before PCD-inducers were added, and were maintained in the same conditions until just before they were chemically fixed (see below). The O₂ concentration remained below 8 p.p.m. throughout the culture period. We shall refer to these conditions as anaerobic. In each experiment, cells were also cultured in parallel in aerobic conditions, in a mixture of 10% CO₂ and 90% air.

We first tested a simian virus 40 (SV40)-transformed human fibroblast cell line that lacks mitochondrial DNA (GM701- ρ^0) and, consequently, is defective in mitochondrial respiration⁷. Like normal cells, they undergo PCD when treated with 1 μ M staurosporine^{8,9}, a broad-spectrum protein-kinase inhibitor¹⁰. We treated the cells with staurosporine (0 to 1 μ M) for 20 hours, either in anaerobic or aerobic conditions, and quantified the resulting PCD by counting the proportion of apoptotic and normal nuclei after staining with bisbenzamide. There was no apparent morphological difference between untreated cells that were grown anaerobically or aerobically, although survival was often lower in anaerobic conditions (Fig. 1a). Cells treated in anaerobic conditions were even more sensitive to staurosporine-induced PCD than cells treated aerobically (Fig. 1a). The cells treated with staurosporine in both conditions displayed nuclear condensation and fragmentation (Fig. 2a-d), although the fragmentation was always greater in anaerobic than in aerobic conditions (Fig. 2b, d). Thus staurosporine-induced PCD seems not to require ROS.

To determine if PCD could occur in anaerobic conditions when induced in other ways, we tested two other inducers: (1) anti-Fas/APO-1 antibodies, which induce PCD by activating cell-surface Fas molecules¹¹; and (2) survival-factor deprivation¹². In the first case, we treated cells of the human lymphoid cell line SKW6.4 with a monoclonal anti-APO-1 antibody¹³. In the second, we treated the interleukin-3 (IL-3)-dependent, murine pro-B cell line BAF3 (ref. 14) with neutralizing, monoclonal anti-IL-3 antibodies¹⁵. In both cases the cells underwent PCD in both anaerobic and aerobic conditions (Fig.

1b, c). Moreover, DNA extracted from BAF3 cells deprived of IL-3 in either anaerobic or aerobic conditions displayed the oligonucleosomal DNA laddering characteristic of PCD (Fig. 2e)¹⁶.

It has been suggested that Bcl-2 inhibits PCD by reducing the generation or effects of ROS (refs 3-5), but when we studied

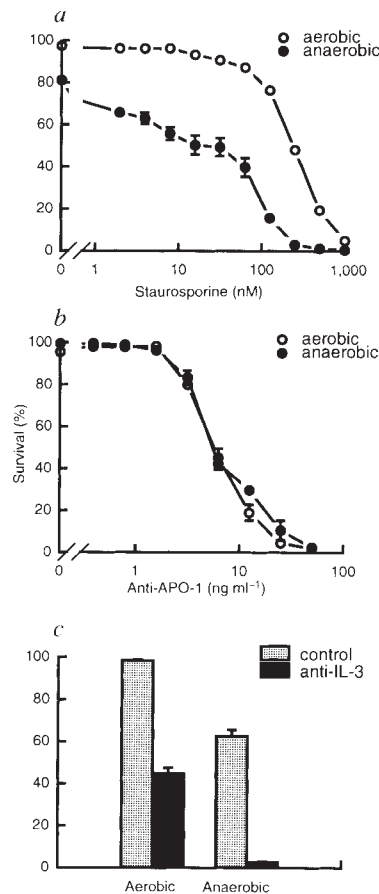


FIG. 1 PCD can occur in both aerobic and anaerobic conditions. *a*, GM701- ρ^0 cells were treated with staurosporine for 20 hours; *b*, SKW6.4 cells were treated with anti-APO-1 for 4 hours; and *c*, BAF3 cells were treated with neutralizing anti-IL-3 monoclonal antibody (hybridoma supernatant diluted 1:20) for 24 hours.

METHODS. GM701- ρ^0 cells were grown in Dulbecco's Modified Eagle Medium containing supplements (DMEM) and 10% fetal calf serum (FCS), as described previously⁹. BAF3/Bo-6 cells¹⁴ (referred to here as BAF3 cells) were grown in the same medium supplemented with 5% conditioned medium from WEHI 3B cells. For each experiment cells were plated onto 96-well tissue-culture plates (Falcon) in 0.1 ml medium at the following densities: GM701- ρ^0 cells, 3×10^3 per well; SKW 6.4 cells, 2×10^4 per well; BAF3 cells, 4×10^3 per well. The GM701- ρ^0 cells were grown overnight before they were transferred to either the anaerobic or the aerobic incubator; all other cells were transferred immediately. All cells were maintained in these conditions for 24 hours before being treated with the test reagents. For anaerobic treatments, cells, media, reagents and equipment were maintained in the anaerobic incubator for at least 24 hours before use. The oxygen concentration in the anaerobic incubator varied between 1.2 and 7.6 p.p.m. After treatment, cells were removed from the incubator, fixed immediately in acetic acid-methanol (1:3) and air dried. Cells were stained with bisbenzamide (Hoechst No. 33258, 0.1 μ g ml⁻¹ in phosphate-buffered saline (PBS)) and percentage survival was determined in an inverted fluorescence microscope by counting apoptotic and non-apoptotic nuclei in each of 6 wells. Data shown are means $\pm$ s.e.m. ($n = 6$) from single representative experiments, which were repeated at least three times with similar results. Monoclonal anti-APO-1 antibodies¹³ were used as supernatants (provided by P. Kramer); monoclonal anti-IL-3 antibodies were used as supernatants prepared from cultures of MP2-8F8.11 hybridoma cells¹⁵.

GM701- ρ^0 cells that were stably transfected with a human *bcl-2* cDNA⁹, they were resistant to staurosporine-induced PCD in both anaerobic and aerobic conditions (Fig. 3a). Similarly, BAF3 cells that were stably transfected with a human *bcl-2* cDNA¹⁴ were protected from PCD in both anaerobic and aerobic conditions when treated with anti-IL-3 antibodies (Figs 2e and 3b). Thus Bcl-2 can protect cells from PCD in ways that do not depend on its inhibiting the generation or effects of ROS.

To check that our anaerobic conditions inhibited the ability of cells to generate ROS, we treated GM701 cells with chemical inducers of ROS: either menadione (which generates superoxide anion and hydrogen peroxide in cells¹⁷), or copper sulphate plus ascorbic acid (which generates hydroxyl radicals in cells¹⁸). Both treatments induced the morphological features of PCD in GM701 cells (Fig. 4a) and GM701- ρ^0 cells (not shown) cultured

in aerobic conditions, but were completely ineffective at doing so in anaerobic conditions (Fig. 4a). Thus our anaerobic conditions were apparently effective in inhibiting the generation of ROS.

Antioxidants have been reported to protect cells against PCD in some circumstances^{3,4,19,27}. As shown in Fig. 4b, in GM701 cells cultured in aerobic conditions, *N*-acetyl-L-cysteine (NAC) and reduced glutathione inhibited menadione-induced PCD, extracellular catalase inhibited Cu(1)-plus-ascorbate-induced PCD, and the glutathione peroxidase mimic Ebselen, in the presence of NAC, inhibited both forms of PCD (Fig. 4b). None of these antioxidants, however, had any protective effect when PCD was induced by either staurosporine (Fig. 4b) or anti-Fas antibodies (in GM701 cells stably transfected with a murine *fas* cDNA, not shown). Furthermore, none of the other antioxidant treatments we tried protected GM701 cells against PCD induced

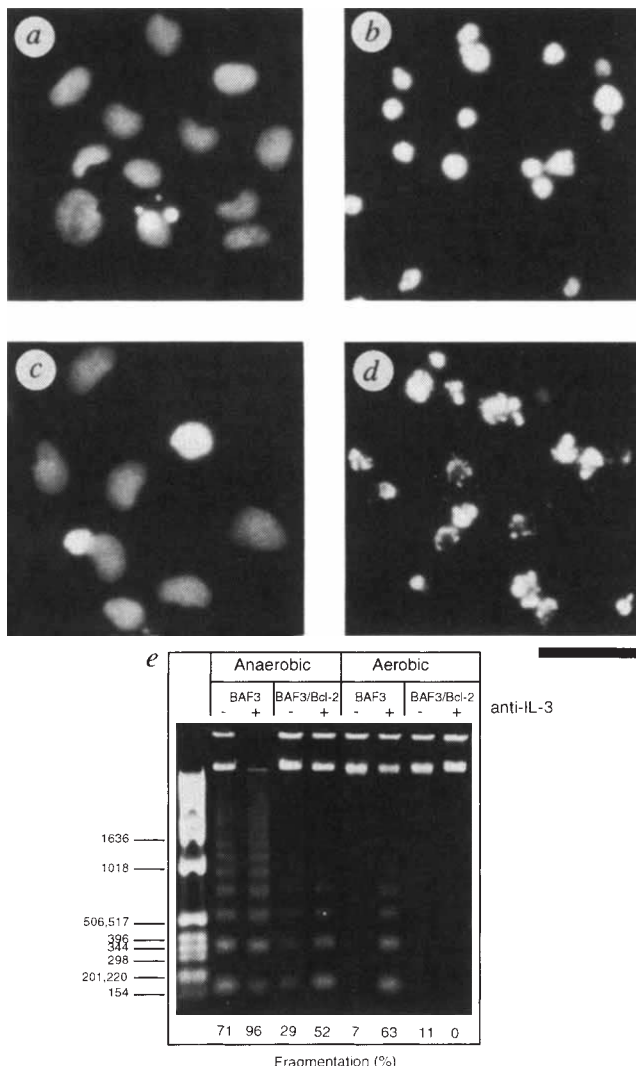


FIG. 3 Effect of staurosporine (a) or anti-IL-3 antibodies (b) on human cells expressing *bcl-2* in aerobic or anaerobic conditions. a, GM701- ρ^0 /Bcl-2 cells were treated with 1 μ M staurosporine for 20 hours. b, BAF3 and BAF3/Bcl-2 cells were treated with anti-IL-3 antibodies for 24 hours. The data in a and b are from the same experiments as in Fig. 1a and c, respectively.

METHODS. Experiments were performed as described in the legend to Fig. 1, except that the cells were maintained in the presence of G418 (Gibco, 1 mg ml⁻¹ for GM701- ρ^0 /Bcl-2 cells, and 2 mg ml⁻¹ for BAF3/Bcl-2 cells). G418 was not present during the experiments.

FIG. 2 Fluorescence micrographs of control or staurosporine-treated GM701- ρ^0 cells, in aerobic (a, b) or anaerobic (c, d) conditions, stained with bisbenzamide, and ethidium bromide-stained DNA (e) extracted from BAF3 and BAF3/Bcl-2 cells treated with or without neutralizing monoclonal anti-IL-3 antibodies in aerobic or anaerobic conditions. In a-d, scale bar, 5 μ m.

METHODS. In a-d, GM701- ρ^0 cells were treated as described in the legend to Fig. 1. Micrographs are from the same experiment as in Fig. 1a, with staurosporine at 1 μ M. In e, 10⁶ BAF3/Bo-6 cells or BAF3/Bo-Bcl-2-15 cells¹⁴ (BAF3/Bcl-2 cells) were cultured in Falcon flasks overnight in either aerobic or anaerobic conditions ([O₂] = 1.2–3.6 p.p.m.). They were then treated with 5% (by volume) anti-IL-3 antibodies or medium for 24 hours, transferred to polypropylene tubes and sealed; for anaerobic cultures, the cells were transferred and sealed in the anaerobic cabinet. DNA was extracted and quantified, and equal amounts were run in each lane of a 1.5% agarose/0.5 $\times$ TBE gel containing ethidium bromide. A video image of the gel was recorded under ultraviolet illumination. The percentage DNA fragmentation was determined by measuring the image intensity of the low (≤ 20 kb)- and high-M, bands, after the background was subtracted, using image analysis software (OptiLab, GRAFTEK) without image manipulation.

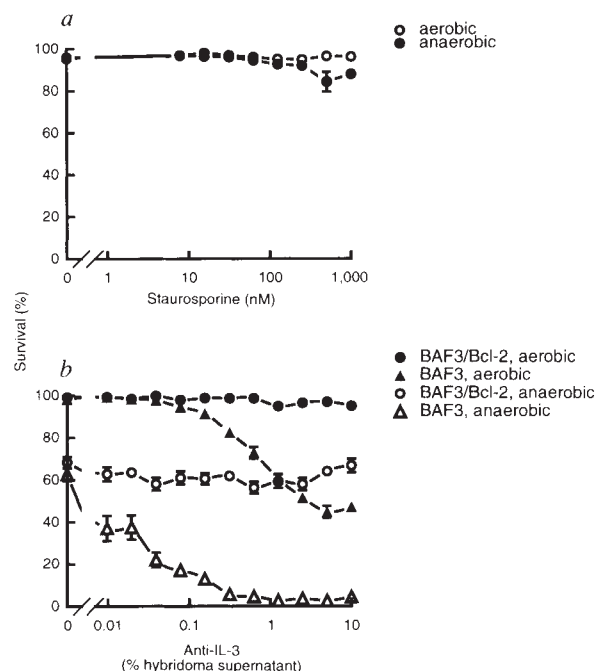
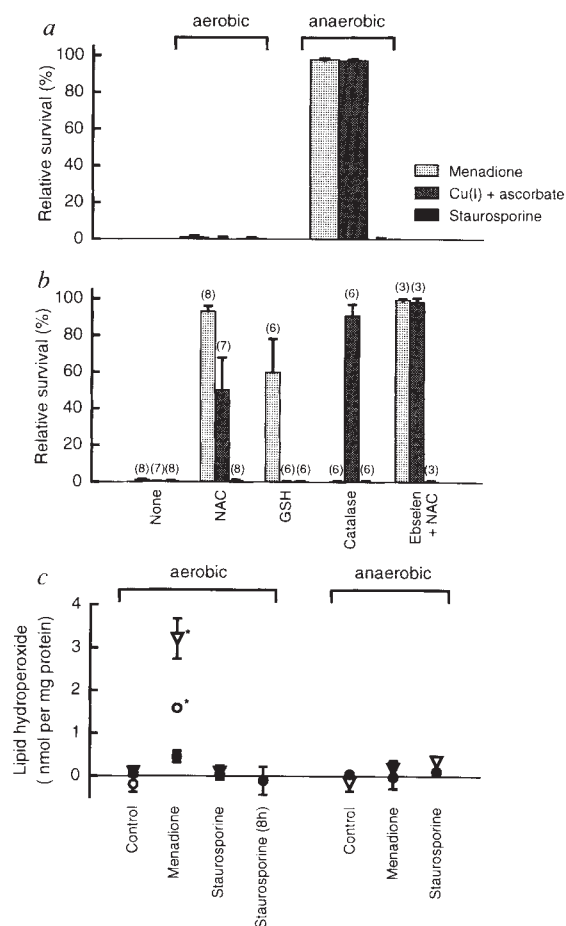


FIG. 4 Effect of anaerobic conditions (a, c) and antioxidants (b) on PCD (a, b) and lipid peroxidation (c) induced in GM701 cells by menadione, Cu(I) + ascorbate, or staurosporine. Final concentrations of drugs in a and b were menadione sodium bisulphite, 200 μ M; copper sulphate/1,10-phenanthroline (1:1), 230 nM; ascorbic acid, 4.5 mM; staurosporine, 1 μ M. In b, antioxidant concentrations were N-acetyl-L-cysteine (NAC), 10 mM; reduced glutathione (GSH), 2 mM; Catalase, 5000 U ml⁻¹; Ebselen, 10 μ M. In c, drug concentrations were menadione, 400 μ M; staurosporine, 2 μ M. Cells were treated with drugs for 18–20 hours in a and b, and 4 hours (or 8 hours where indicated) in c. In c, asterisks indicate, with $\geq 98\%$ confidence, that the means are greater than 0 (one-tailed t-test with Bonferroni correction, $\nu=4$ or 6, 16 means). All other effects were not statistically significant ($p>0.05$). METHODS. In a and b, cultures were equilibrated and treated as in Fig. 1. Antioxidants (b) were added 30 to 60 min before the addition of PCD inducers. Survival in the presence of the PCD-inducer was normalized to survival in the absence of the inducer (and in presence of the antioxidants in b), which was always greater than 90%. A representative experiment is shown in a; this experiment was repeated 3 times with similar results. In b, data are shown as means $\pm$ s.e.m. of the number of experiments indicated in parentheses. In c, confluent cultures of GM701 cells (about 2×10^7 per treatment) were treated as above, the cells and culture medium were collected and sealed in 50-ml centrifuge tubes and chilled on ice. The cells were centrifuged at 4 $^{\circ}$ C, washed 3 times with ice-cold PBS, and assayed for lipid peroxidation by the FOX-2 method²⁹ and normalized to total protein, which was determined by the BCA assay (Pierce, USA). Data in c are the means $\pm$ s.e.m. ($n=3$ or 4) for each experiment.



by staurosporine, anti-Fas antibodies, or the two oxidant treatments (Fig. 4). These included the free-radical scavengers BHA (200 μ M), BHT (200 μ M), NDGA (25 μ M), α -tocopherol acetate (230 μ M), *N*-propyl gallate (100 μ M) and Trolox (200 μ M), the electron-spin trap 4-hydroxy-Tempo (3 mM), the thiol carbamate PDTC (100 μ M), the iron chelator deferoxamine (1 mM), and the superoxide dismutase mimic Euk-8 (C7, 40 μ M)²⁸, with or without catalase. Thus, although antioxidants have been shown to block certain ROS-dependent signal-transduction pathways that induce PCD^{19, 21, 27}, they apparently do not block the effector phase of PCD.

Although lipid peroxidation, a measure of oxidative damage to membrane lipids, has been reported to increase in cells undergoing PCD^{3, 4}, we found no evidence for this in stauro-

sporine-treated GM701 cells grown either aerobically or anaerobically (Fig. 4c). As expected, menadione induced lipid peroxidation in aerobic, but not anaerobic, conditions (Fig. 4c).

The low oxygen pressures in our anaerobic conditions (<8 p.p.m.) should exclude the formation of most, if not all, ROS, although they would not exclude oxygen-independent redox or electron transfer reactions. Our findings therefore suggest that ROS are not an essential part of PCD: although ROS can activate the death programme, they seem not to be required to execute it. Moreover, our results suggest that Bcl-2 protects cells against PCD in ways that do not depend on the protein's ability to inhibit the production or effects of ROS. It is possible that ROS are generated as a consequence of PCD and that Bcl-2 inhibits their generation, their effects, or both, simply by preventing or delaying cell death. □

Received 7 December 1994; accepted 17 March 1995.

- Dexter, T. M., Raff, M. C. & Wyllie, A. H. (eds) *Phil. Trans. R. Soc. B* **345**, 231–233 (1994).
- Reed, J. C. *J. Cell Biol.* **124**, 1–6 (1994).
- Hockenbery, D. M., Oltvai, Z. N., Yin, X.-M., Millman, C. L. & Korsmeyer, S. J. *Cell* **75**, 241–251 (1993).
- Kane, D. J. *et al. Science* **262**, 1274–1277 (1993).
- Veis, D. J., Sorenson, C. M., Shutter, J. R. & Korsmeyer, S. J. *Cell* **75**, 229–240 (1993).
- Buttke, T. M. & Sandstrom, P. A. *Immun. Today* **15**, 7–10 (1994).
- King, M. P. & Attardi, G. *Science* **246**, 500–503 (1989).
- Jacobson, M. D., Burne, J. F. & Raff, M. C. *EMBO J.* **13**, 1899–1910 (1994).
- Jacobson, M. D. *et al. Nature* **361**, 365–369 (1993).
- Tamaoki, T. & Nakano, H. *BioTechnology* **8**, 732–735 (1990).
- Nagata, S. & Golstein, P. *Science* **267**, 1449–1456 (1995).
- Raff, M. C. *Nature* **356**, 397–400 (1992).
- Trauth, B. E. *et al. Science* **245**, 301–305 (1989).
- Marvel, J., Perkins, G. R., Lopez-Rivas, A. & Collins, M. K. L. *Oncogene* **9**, 1117–1122 (1994).
- Abrams, J. S. & Pearce, M. K. *J. Immun.* **140**, 131–137 (1988).
- Wyllie, A. H., Morris, R. G., Smith, A. L. & Dunlop, D. J. *Path.* **142**, 67–77 (1984).
- Cadenas, E. A. *Rev. Biochem.* **58**, 79–110 (1989).

- Halliwell, B. & Gutteridge, J. M. C. *Meth. Enzym.* **186**, 1–85 (1990).
- Ratan, R. R., Murphy, T. H. & Baraban, J. M. *J. Neurosci.* **14**, 4385–4392 (1994).
- Behl, C., Davis, J. B., Lesley, R. & Schubert, D. *Cell* **77**, 817–827 (1994).
- Franklin, J. L., Miller, T. M. & Johnson, E. M. *Soc. Neurosci. Abstr.* **20**, 432 (1994).
- Lennon, S., Martin, S. & Cotter, T. *Cell Prolif.* **24**, 203–214 (1991).
- Wolfe, J. T., Ross, D. & Cohen, G. M. *FEBS Lett.* **352**, 58–62 (1994).
- Abello, P. A., Fidler, S. A., Bulkley, G. B. & Buchman, T. G. *Archs. Surg.* **129**, 134–140 (1994).
- Sandstrom, P. A., Mannie, M. D. & Buttke, T. M. *J. Leukoc. Biol.* **55**, 221–226 (1994).
- Ratan, R. R., Murphy, T. H. & Baraban, J. M. *J. Neurochem.* **62**, 376–379 (1994).
- Schulze-Osthoff, K., Krammer, P. H. & Dröge, W. *EMBO J.* **13**, 4587–4596 (1994).
- Baudry, M. *et al. Biochem. biophys. Res. Commun.* **192**, 964–968 (1993).
- Nourooz-Zadeh, J., Tajaddini-Sarmadi, J. & Wolff, S. P. *Analyt. Biochem.* **220**, 403–409 (1994).

ACKNOWLEDGEMENTS. We thank T. Mitchison for suggesting the anaerobic experiments; R. Cramack for use of the anaerobic cabinet; J. Abrams (DNAX), M. Collins, P. Golstein, P. Krammer, B. Malfoy and S. Nagata for cells and reagents; A. White and S. Wolff for technical assistance; B. Halliwell, A. McLean and S. Wolff for helpful advice and comments; and Y. Tsujimoto, G. Cohen and their colleagues for communicating results before publication. M.D.J. is supported by a British Telecom-sponsored grant from the Multiple Sclerosis Society (UK).

A kinase from fission yeast responsible for blocking mitosis in S phase

Hiroshi Murakami & Hiroto Okayama

Department of Biochemistry, Faculty of Medicine, The University of Tokyo, Bunkyo-ku, Tokyo, Japan; The Okayama Cell Switching Project, ERATO, JRDC, Sakyo-ku, Kyoto, Japan

In virtually all eukaryotes, mitosis starts after the completion of DNA synthesis. This orderly process is ensured by the checkpoint mechanism that blocks the onset of mitosis while DNA is being synthesized or is damaged. In the fission yeast *Schizosaccharomyces pombe*, this mechanism involves some *rad*⁺ and *hus*⁺ genes¹⁻⁴. However, it is not known how the checkpoint system monitors these events. Recently a multicopy suppressor of a temperature-sensitive DNA polymerase- α mutant was isolated. This gene, named *cds1*⁺ (checking DNA synthesis), encodes a typical protein kinase. Here we report that this protein kinase is a key component of the DNA replication-monitoring S/G2 checkpoint system. Our data suggest that its primary role is to monitor DNA synthesis by interacting with DNA polymerase α and send a signal to block the onset of mitosis while DNA synthesis is in progress.

We generated several temperature-sensitive cell division cycle (*cdc*) mutants of *S. pombe* and searched for those that strongly interacted with the existing checkpoint-deficient mutants. One of these, *swi7-H4*, was defective in S-phase progression and possessed such a phenotype. It was synthetically lethal with *rad3-136* (ref. 2), and its ability for growth was significantly impaired when combined with *rad1-1* or *rad9-192* (refs 1 and 2). Moreover, it produced abnormal mitotic cells at the non-permissive temperature (see below). The DNA polymerase α gene (*pola*⁺)⁵ fully complemented *swi7-H4*, even at 36 °C, which is 2 °C above the restriction temperature. Crossing and random spore analysis revealed that the locus of this mutation was completely linked to *swi7*⁺ (*pola*⁺)⁶. Thus the mutated gene was *pola*⁺, which was involved in checkpoint function.

We were therefore interested in identifying molecules that interact with *Pola*. The mutant *swi7-H4* was transfected with *S. pombe* complementary DNA and genomic libraries, and two distinct suppressive clones were isolated: *pola*⁺ and *cds1*⁺. The suppression by *cds1*⁺ was seen up to 36 °C. The *cds1*⁺ cDNA is 1.5 kilobases (kb) long and encodes a 460 amino-acid protein (estimated relative molecular mass (*M_r*) of 51,870) with a complete set of serine/threonine protein kinase domains in its C-terminal region (Fig. 1). This kinase has the highest homology with SPK1 of *Saccharomyces cerevisiae*⁷ (35% amino-acid identities) but differs in several respects. SPK1 has a long C-terminal extension, is involved in general checkpoint function, and is essential for viability and growth⁸⁻¹⁰. By contrast, *Cds1* has no C-terminal extension (Fig. 1), is not required for growth and viability, and is involved in S-phase-specific checkpoint function as described below.

By gene replacement of the coding region with *ura4*⁺, *cds1*⁻ disruptants were constructed which were viable with morphology and growth rate indistinguishable from those of the wild-type cells. How-

ever, they displayed profound sensitivity to hydroxyl urea (HU, a DNA synthesis inhibitor) but not ultraviolet light (UV) (Fig. 2a, b), phenotypically similar to the recently reported mutants¹¹. Their HU sensitivity was completely suppressed by *cds1*⁺ but not by *pola*⁺, *rad1*⁺ (ref. 11) (Fig. 2e) or other checkpoint-related genes including *rad9*⁺ (ref. 12) and *chk1*⁺ (ref. 13) (data not shown).

In *S. pombe*, Cdc25 phosphatase is required for the activation of Cdc2 kinase, and so its temperature-sensitive mutant is unable to start mitosis at the restrictive temperature¹⁴. If the HU sensitivity of *cds1*⁻ disruptants were a result of a defect in the S/G2 checkpoint function, it would be rescued by the inactivation of *cdc25*⁺: this was indeed the case. The temperature-sensitive *cdc25-22 cds1*⁻ double mutant grew as rapidly as the *cdc25-22* cells. When the double mutant was exposed to HU and shifted to the non-permissive temperature at various times, the further loss of its viability was markedly suppressed (Fig. 2f). Under these conditions, most of the cells accumulated in early S within the initial exposure to HU of 2-3 h. Taken together these results indicate that *cds1*⁺ is involved in checkpoint function.

Consistent with this putative role, *cds1*⁺ rescued nearly completely the HU but not the UV sensitivity of the *rad1* (Fig. 2c, d), *rad3* and *rad9* mutants (data not shown) with concomitant suppression of 'cut' that emerged during HU treatment and its removal. This abnormal mitosis, characterized by nuclear divi-

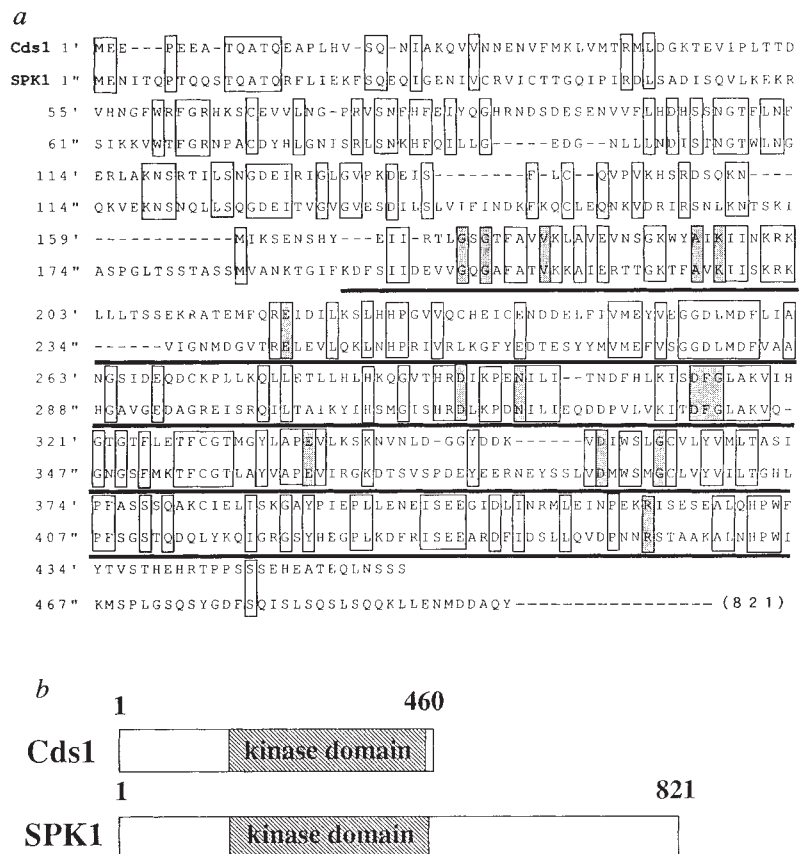


FIG. 1 a, The amino-acid sequences of Cds1 and SPK1 (ref. 7). The amino-acid residues identical between Cds1 and SPK1 are boxed, kinase domains are underlined, and highly conserved amino acids in the domains are shaded. b, Schematic representation of the structure of Cds1 and SPK1.

METHODS. The *h*⁻ *leu1-32 swi7-H4* cells were transfected with a *S. pombe* cDNA library²¹ or genomic libraries containing autonomous replicating sequences (ars) and the *LEU2* gene (K. Okazaki and H.O., unpublished data), and incubated on MMA plates first at 25 °C for 24 h and then at 34 °C for 3-4 days. Complementing plasmids were recovered in *Escherichia coli*²¹. The 1.5 kb cDNA and the 3.0 kb genomic *Spel*/*Pst*I fragment that had complementation activity were sequenced by the dideoxynucleotide method²². Sequence data for the *cds1*⁺ cDNA have been submitted to EMBL, accession number X85040.

FIG. 2 *a*, HU and *b*, UV sensitivity of wild-type (open diamonds), *cdc2-3w* (open circles), *rad1-1* (open triangles) and *cds1*⁻ (filled squares) cells. *c*, Suppression of HU. *d*, suppression of UV sensitivity of *rad1-1* cells by pCL-*cds1*⁺ (filled diamonds), pCL-X (empty vector) (open circles) or pAL-*rad1*⁺ (open squares). *e*, *f*, Suppression of HU sensitivity of *cds1*⁻ cells by (e) pAL-*cds1*⁺ (filled squares), pAL-X (open diamonds), pAL-*polα*⁺ (open circles) and pAL-*rad1*⁺ (open triangles) or by (f) inactivation of *cdc25*. The *cdc25-22* and *cdc25-22 cds1*⁻ cells were shifted from 25 °C to 36 °C after 2 h (diamonds), 3 h (circles) or 4 h (triangles) HU exposure or none (squares), and further incubated for the indicated times. METHODS. The *SphI-EcoRV* fragment of *cds1*⁺ (75% of open reading frame) was replaced with *ura4*⁺, and the resulting replaced *PstI-SpeI* fragment was transfected into a diploid strain to construct disruptants²¹. The *h*^{-S} *cds1::ura4*⁺ *ura4D-18 leu1-32* (*cds1*⁻ disruptant), *h*^{-S} *leu1-32* (wild type), *h*^{-S} *cdc2-3w leu1-32*, and *h*^{-S} *rad1-1 leu1-32* cells were grown to a mid-log phase ($5-8 \times 10^6$ ml⁻¹) in yeast extract media at 30 °C. For Fig. 2*a*, HU was added to the culture to final 12 mM at time 0. Portions were taken, plated on yeast extract and cell viability was determined. For Fig. 2*b*, cells in a mid-log phase were plated and exposed to UV at a dose of 3 J m⁻². Unirradiated plates provided negative control. For Fig. 2*c-e*, *rad1*⁻ (*c*, *d*) and *cds1*⁻ (*e*) cells were transformed with *rad1*⁺, *cds1*⁺ or none (X) inserted in the pCL or pAL yeast expression vector²¹. Minimum medium was used. For Fig. 2*f*, the *h*^{-S} *cdc25-22 cds1*⁻ *leu1-32* and *h*^{-S} *cdc25-22 leu1-32* cells were used.

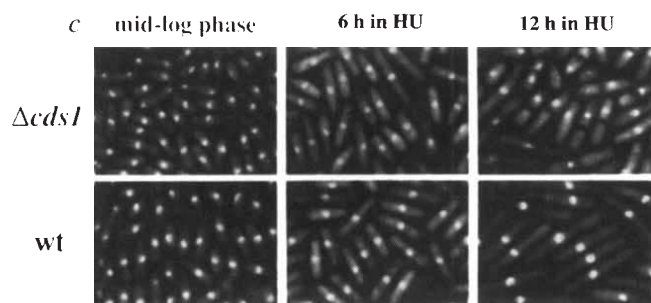
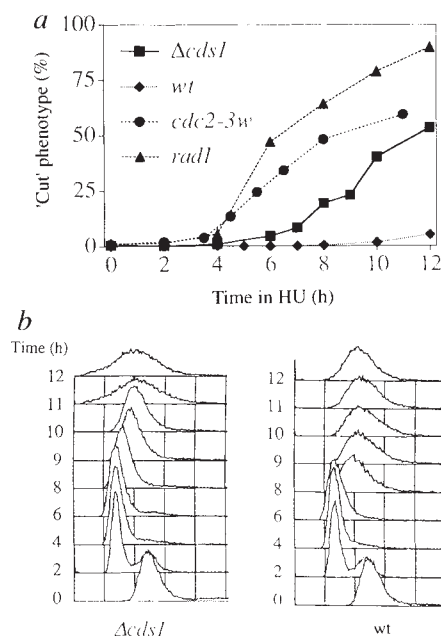
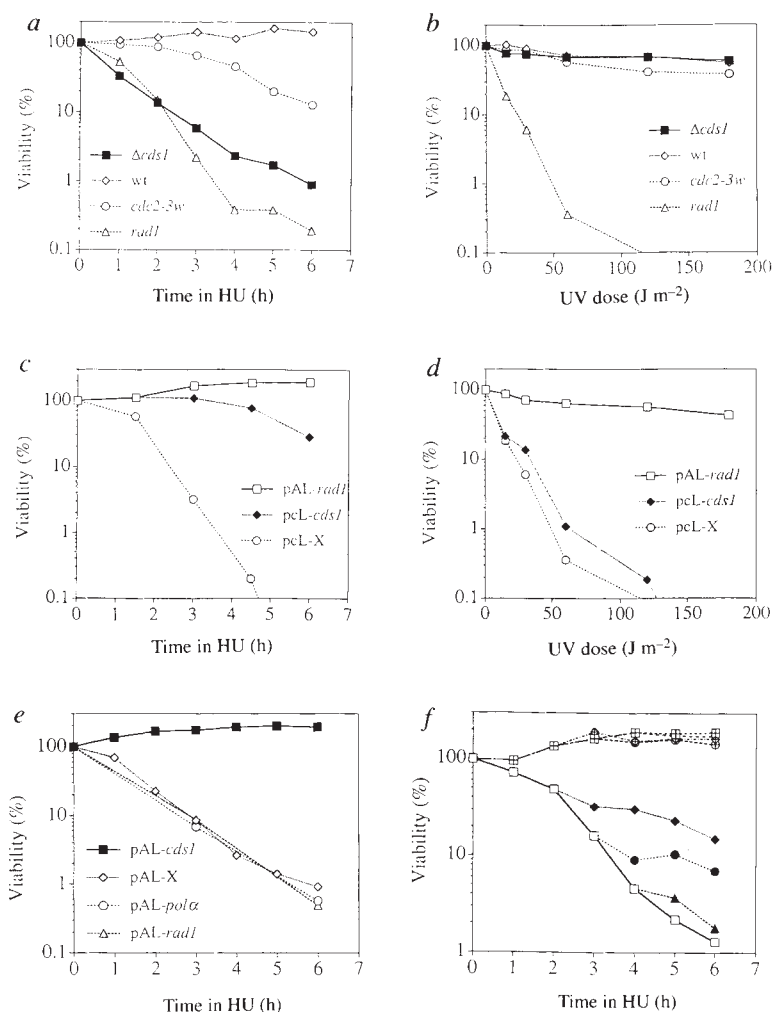


FIG. 3 HU arrests *cds1*⁻ cells in early S, but aberrant mitosis occurs when DNA synthesis begins. *a*, Population of 'cut' in *cds1*⁻ disruptant (squares), wild-type (wt) (diamonds), *cdc2-3w* (circles) and *rad1-1* (triangles) during HU exposure. *b*, DNA content of *cds1*⁻ disruptant and wild-type cells analysed by flow cytometry. *c*, The 4,6-diamino-2-phenylindole (DAPI)-stained *cds1*⁻ disruptant and wild-type cells under various conditions. METHODS. The *h*^{-S} *cds1::ura4*⁺ *ura4D-18 leu1-32*, *h*^{-S} *leu1-32*, *h*^{-S} *cdc2-3w leu1-32*, and *h*^{-S} *rad1-1 leu1-32* cells were grown to a mid-log phase in yeast extract medium at 25 °C. HU was added to the culture to 12 mM at time 0. DAPI staining and flow cytometry analysis were performed with cells fixed in 70% (by volume) ethanol, as previously described²¹.

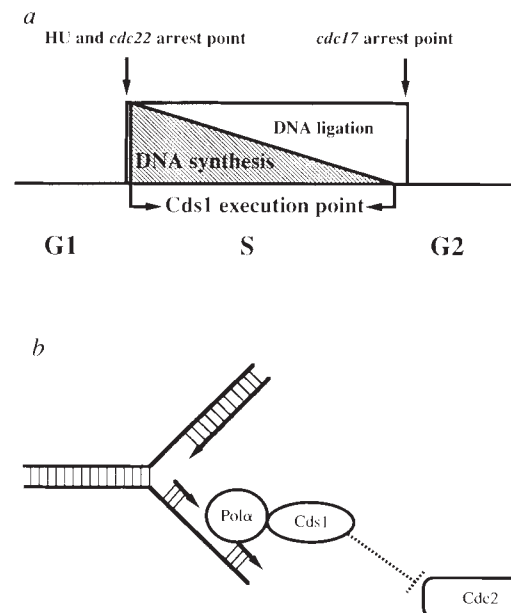


FIG. 4 Schematic representation of (a) proposed Cds1 execution point and (b) Cds1 function.

sion by septum and uneven cell division generating aneuploid and anucleate cells, is frequently seen in the cells defective in checkpoint function^{3,15,16}. Thus *cds1*⁺ plays an important role in the DNA-monitoring checkpoint mechanism.

When exposed to HU, the *rad1-1* and *cdc2-3w* cells started to display 'cut' at around 4 h without DNA replication^{1,3,5} (Fig. 3a). In contrast, HU-exposed *cds1*⁻ cells did not produce 'cut' until after 6 h. When continuously exposed to HU, the wild-type cells spontaneously initiated DNA replication at 7–8 h (ref. 17) with no 'cut' (Fig. 3). The *cds1*⁻ cells similarly initiated DNA synthesis after 8 h, but with rapid production of 'cut'. The timing of DNA synthesis and the emergence of 'cut' coincided even when HU was removed after 3–6 h treatment (data not shown). Thus the checkpoint defect of *cds1*⁻ cells was most evident during DNA synthesis. Consistently, in 3 mM HU, which allows but slightly inhibits DNA synthesis, a much higher population of the *cds1*⁻ cells (90% after a 24 h incubation) displayed 'cut', whereas the wild-type cells had no 'cut' despite a twofold increase in generation time.

These results were confirmed with the temperature-sensitive *cdc22-M45 cds1*⁻ double mutant. The large subunit of ribonucleotide reductase, a putative target for HU action¹⁸, is encoded by *cdc22*⁺. The double mutant at non- or semi-permissive temperatures mimics HU-treated *cds1*⁻ cells. At 25 °C it grew with little 'cut' (5%) and, upon shift to 36 °C, was arrested in early S with no increase in 'cut' until after 4 h. But 'cut' drastically increased thereafter and reached 40% after 8 h (6% for *cdc22-M45*). However, when the mutant was exposed to a semi-permissive temperature (30 °C), 'cut' immediately began to accumulate

and finally reached >50% with little cell growth. The close association of the *cds1*⁺'s checkpoint function with DNA synthesis was further confirmed with the *swi7-H4 cds1*⁻ double mutant. At the permissive temperature the double mutant was viable with a reduced growth rate and little 'cut'. But after 8 h exposure to the non-permissive temperature, 'cut' accumulated to 30% (15% for *swi7-H4*).

These observations led us to examine whether *cds1*⁺ was also required for checkpoint block after the completion of DNA synthesis but not before the end of the S phase. Cells lacking DNA ligase mimic such a situation. In *S. pombe*, *cdc17*⁺ encodes DNA ligase¹⁹. The *cds1*⁻ *cdc17-K42* double mutant, however, had no detectable checkpoint defect at any temperature tested (data not shown).

From these data we conclude that *cds1*⁺ executes its checkpoint function while DNA synthesis is in progress (Fig. 4a). Its strong interaction with Polα and its requirement for checkpoint block of mitosis during DNA polymerization suggest that the primary role of *cds1*⁺ is to monitor DNA replication via Polα (Fig. 4b). Our data, however, do not rule out the possibility that it might have an additional role as an auxiliary factor for DNA polymerization. How Cds1 senses the status of DNA synthesis from Polα, and where the signal is transmitted, is unclear at present. However, it is likely to do this by binding directly to Polα, as it effectively rescues the temperature-sensitive *polα* mutant. A protein kinase of a similar size associated with mammalian Polα has recently been reported²⁰. Our results also indicate that there must be at least two more distinct sensors, which monitor pre- and post-DNA synthesis states for checkpoint arrest. □

Received 17 January; accepted 2 March 1995.

- Rowley, R., Subramani, S. & Young, P. G. *EMBO J.* **11**, 1335–1342 (1992).
- Al-Khodairy, F. & Carr, A. M. *EMBO J.* **11**, 1343–1350 (1992).
- Enoch, T., Carr, A. M. & Nurse, P. *Genes Dev.* **6**, 2035–2046 (1992).
- Al-Khodairy et al. *Molec. Biol. Cell* **5**, 147–160 (1994).
- Damagnez, V., Tillit, J., deRecondo, A. M. & Baldacci, G. *Molec. Gen. Genet.* **226**, 182–189 (1991).
- Singh, J. & Klar, A. J. S. *Nature* **361**, 271–273 (1993).
- Stern, D. F. P., Zheng, P., Beidler, D. R., & Zerillo, C. *Molec. cell. Biol.* **11**, 987–1001 (1991).
- Zheng, P. et al. *Molec. cell. Biol.* **9**, 5829–5842 (1993).
- Weinert, T. A., Kiser, G. L. & Hartwell, L. H. *Genes Dev.* **8**, 652–665 (1994).
- Allen, J. B., Zhou, Z., Siede, W., Friedberg, E. C. & Elledge, S. J. *Genes Dev.* **8**, 2416–2428 (1994).
- Sunnerhagen, P., Seatron, B. L., Nasim, A. & Subramani, S. *Molec. cell. Biol.* **10**, 3750–3760 (1990).
- Murray, J. M., Carr, A. M., Lehmann, A. R. & Watts, F. Z. *Nucleic Acids Res.* **19**, 3525–3541 (1991).

- Walworth, N., Davey, S. & Beach, D. *Nature* **363**, 368–371 (1993).
- Fantes, P. *Nature* **279**, 428–430 (1979).
- Enoch, T. & Nurse, P. *Cell* **60**, 665–673 (1990).
- Hirano, T., Funahashi, S., Uemura, T. & Yanagida, M. *EMBO J.* **5**, 2973–2979 (1986).
- Sazer, S. & Nurse, P. *EMBO J.* **13**, 606–615 (1994).
- Sarabia, M.-J. F., McInerney, C., Harris, P., Gordon, C. & Fantes, P. *Molec. Gen. Genet.* **238**, 241–251 (1993).
- Nasmyth, K. A. *Cell* **12**, 101–106 (1977).
- Peck, V. M., Gerner, E. W. & Cress, A. E. *Biochem. biophys. Res. Commun.* **190**, 325–331 (1993).
- Tanaka, K. et al. *EMBO J.* **11**, 4923–4932 (1992).
- Sanger, F., Nicklen, S., Barrell, B. G., Smith, A. J. H. & Roe, B. A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).

ACKNOWLEDGEMENTS. We thank J. Singh and A. J. S. Klar for a strain with a *ura4*⁺ marker integrated near the *swi7*⁺ gene. We also thank K. Okazaki for providing *S. pombe* genomic libraries. This work was supported by Grants-in-aid for Scientific Research from the Ministry of Education, Science and Culture of Japan, and grants from the Asahi Glass Foundation.

Connecting a promoter-bound protein to TBP bypasses the need for a transcriptional activation domain

Sukalyan Chatterjee & Kevin Struhl*

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA

* To whom correspondence should be addressed

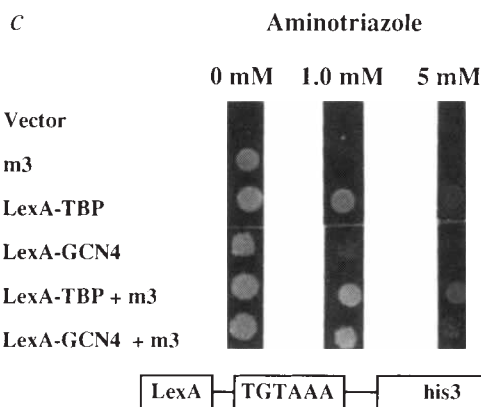
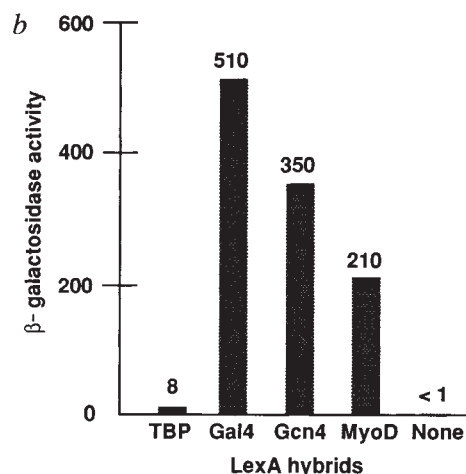
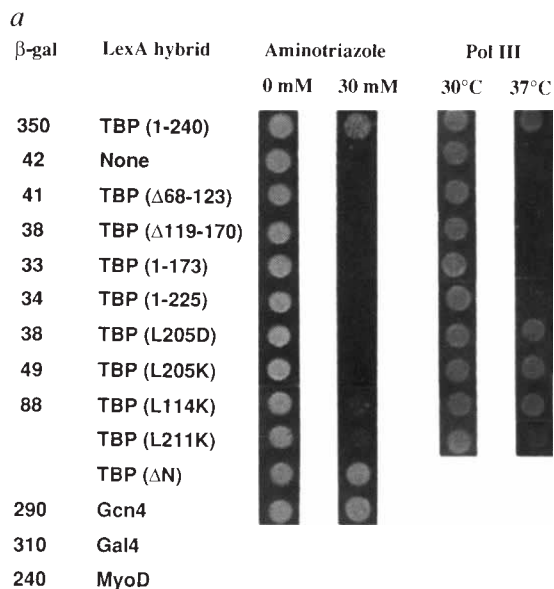
BIOCHEMICAL analyses have suggested potential targets for transcriptional activation domains, which include several components of the RNA polymerase II machinery¹⁻⁷, as well as the chromatin template⁸⁻¹². Here we examine the mechanism of transcriptional activation in yeast cells by connecting a heterologous DNA-binding domain (LexA) to the TATA-binding protein (TBP). LexA-TBP efficiently activates transcription from a promoter containing a LexA operator upstream of a TATA element. Activation is promoter-specific and is sensitive to mutations on the DNA-binding surface of TBP; hence it is not due to a fortuitous activation domain on TBP. Thus a promoter-bound protein lacking an activation domain can stimulate transcription if it is directly connected to TBP. This suggests that recruitment of TBP to the promoter can be a rate-limiting step for transcription *in vivo*, and that interactions between activation domains and factors that function after TBP recruitment can be bypassed for activation.

LexA-TBP activates transcription from a promoter containing a single LexA operator 45 base pairs (bp) upstream of the *his3* TATA element and structural gene (Fig. 1a). Activation by LexA-TBP requires an intact TBP core domain; C-terminal and internal deletions at widely separated regions of the core domain are non-functional. LexA-TBP behaves comparably to LexA hybrids containing strong activation domains from yeast and mammalian proteins (Gal4, Gcn4, MyoD).

Because a wide variety of sequences can activate transcription in yeast^{13,14}, it was critical to exclude the possibility that TBP contains a sequence(s) that fortuitously behaves as a conventional activation domain. Such activation domains are typically localized to short regions of a protein, and they are remarkably insensitive to amino-acid substitutions or small deletions^{13,15,16}.

FIG. 1 Transcriptional activation by LexA-TBP and its derivatives. a, Strains containing the indicated LexA-TBP derivative and a promoter containing a LexA operator 45 bp upstream of the *his3* TATA element and structural gene²³ were tested for growth in aminotriazole (AT). The degree of AT resistance is directly related to the level of *his3* transcription^{24,25}. A control strain containing a comparable promoter lacking the LexA operator does not grow in 1 mM AT in the presence or absence of LexA-TBP. Expression (β -gal) was directly measured with a *LacZ* derivative of the promoter on the *URA3* multicopy plasmid YEp356 (ref. 26). LexA-TBP derivatives were also tested for their ability to support RNA polymerase III (pol III) transcription by complementation of a temperature-sensitive, pol III-specific derivative of TBP (F155S) at 37 (ref. 27). b, Similar experiment except that strains contain YEp21-Sc3423 (ref. 13), a *LEU2* multicopy plasmid with the LexA operator located 81 and 135 bp upstream of the *cyc1* TATA elements and *lacZ* structural gene. c, Similar experiment except that the promoter differs by a single bp in the TATA element (TGTAAG), and strains are tested at lower AT concentrations. Where indicated, strains also contain a *URA3* centromeric plasmid expressing TBP^{m3}, a derivative that binds with increased affinity to TGTAAG²¹.

METHODS. The centromeric *TRP1* expression vector YCp91 (ref. 23) containing LexA-TBP was generated by replacing the region encoding the HA1 epitope and the *CYC8* structural gene of YCp91-LexA-CYC8



(ref. 23) with the TBP structural gene²⁸. The resulting LexA-TBP hybrid contains 12 residues encoding the SV40 nuclear localization signal between the full-length LexA (1-202) and TBP (1-240) moieties. Deleted versions of LexA-TBP were generated by cleaving at the relevant *Pst*I, *Bgl*II, *Xba*I and *Hind*III sites. TBP derivatives lacking the non-conserved and non-essential N-terminal region²⁸ or containing amino-acid substitutions at residues on the DNA-binding surface of TBP^{19,20} were converted into their LexA-TBP counterparts by replacing the relevant fragments. LexA-Gcn4 (ref. 13), LexA-Gal4 (ref. 29) and LexA-MyoD (ref. 30) have been described previously. The activities of β -galactosidase (average of 3-5 independent transformants) are normalized to the OD₆₀₀ of cells at the time of collection and are accurate to $\pm 20\%$. All yeast strains were derived from FT4 (ref. 23).

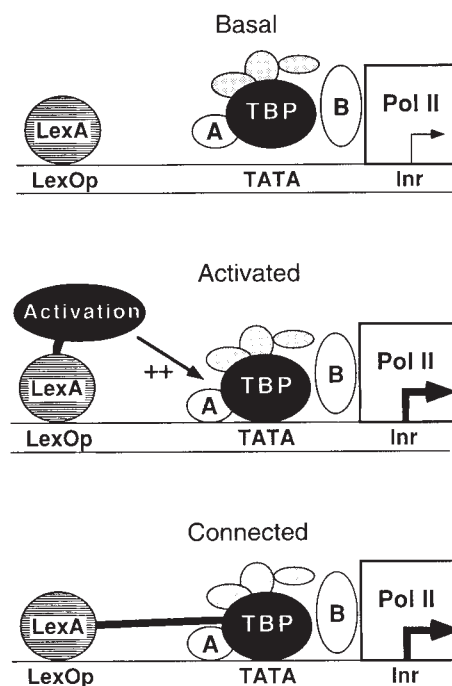


FIG. 2 Connecting a promoter-bound protein to TBP overrides the need for a transcriptional activation domain. Interactions of LexA hybrid proteins, TBP, TBP-associated factors (TAFs; shaded ovals), TFIIA, TFIIB and RNA polymerase II at a promoter containing a LexA operator, TATA element and initiator (Inr). In basal transcription (LexA alone), the interaction between TBP and the TATA element is limiting (indicated by space between the two components), thereby leading to low levels of mRNA (thin arrow). Activation domains increase the recruitment or stability of TBP to the TATA element, thereby stimulating transcription (thick arrow). The target of the activation domain is not specified in the illustration (see text for discussion). A similar increase in transcription occurs when LexA is directly connected (thick line) to TBP. See text for details and limitations of this model.

We therefore examined LexA–TBP derivatives with single mutations at various positions on the DNA-binding surface of TBP^{17,18} that affect TATA-element binding *in vitro*^{19,20}. Such mutations are very unlikely to affect TBP structure or the activity of a conventional activation domain. However, if activation requires that a single LexA–TBP molecule interact both with a LexA operator and the adjacent TATA element, such mutated derivatives should be functionally impaired. Indeed, four such LexA–TBP derivatives show significantly reduced or non-detectable levels of expression (Fig. 1a). Moreover, these proteins are properly folded and otherwise functional *in vivo*, as indicated by their ability to support RNA polymerase III transcription. Thus activation by LexA–TBP requires binding of the TBP moiety to the TATA element, and it is highly sensitive to amino-acid substitutions at various locations on the DNA-binding surface.

Two additional observations provide independent evidence that LexA–TBP behaves differently from a conventional activator. First, LexA–TBP is inactive on a promoter containing a LexA operator upstream of the *cycl* TATA elements, whereas LexA hybrids containing Gcn4, Gal4 or MyoD activation domains function efficiently (Fig. 1b). Thus, unlike a conventional activator, stimulation by LexA–TBP is strongly influenced by promoter context. Second, when tested on a promoter with a LexA operator upstream of a mutated *his3* TATA element (TGTAAG), LexA–TBP activates transcription to a modest extent whereas LexA–GCN4 activates very weakly (Fig. 1c). Introduction of TBP^{m3}, a derivative with increased affinity for TGTAAG sequences²¹, does not significantly affect activation by LexA–TBP, whereas it clearly increases activation by LexA–GCN4 (and the native GCN4 and Gal4 activators²¹). Thus transcription by LexA–TBP, unlike LexA–GCN4, is not dependent on a physically unconnected molecule bound to the TATA element.

Taken together, the above results indicate that activation by LexA–TBP is due to the physical connection between the LexA and TBP domains that allows binding of a single molecule to the adjacent LexA operator and TATA element (Fig. 2). Because LexA and the physical connection are very unlikely to affect inherent TBP functions, activation by LexA–TBP almost certainly involves increased interaction of the TBP moiety with the TATA element as a result of its physical connection to a protein

bound to a nearby site. This effect could be solely due to cooperative binding by physically connected domains, but it may also involve a local disruption of chromatin caused by LexA binding to its operator which increases binding of TBP to the TATA element. In either case, the results strongly suggest that interaction of TBP with the TATA element can be a rate-limiting step for transcription in yeast cells, and that activation domains can increase recruitment (or stabilize the interaction) of TBP to the promoter (Fig. 2). Independent evidence for this crucial role of activation domains has been obtained by measuring the rate at which TBP can productively access the chromatin template *in vivo*²². Although the mechanism for increased recruitment of TBP is unknown, an attractive hypothesis is that the artificial LexA–TBP connection mimics the interaction that normally occurs between the activation domain and TBP (or associated proteins).

LexA–TBP does not contain an activation domain, yet it stimulates transcription to a level comparable to native yeast activator proteins. This indicates that it is possible to generate circumstances in which interactions between activation domains and components of the basic machinery that function after TBP binding are not essential for transcriptional activation *in vivo*. However, our results do not preclude the importance of such interactions at native promoters *in vivo*. Such interactions are likely to be important for achieving maximal levels of transcription, particularly at promoters where the TBP–TATA element interaction is not the major limiting step.

Note added in proof: Our observation that LexA–TBP does not stimulate the *CYC1* promoter is consistent with the recent finding that TBP is bound to *CYC1* TATA elements *in vivo* in the absence of upstream activator proteins³¹. □

Received 19 September 1994; accepted 2 March 1995.

- Ingles, C. J., Shales, M., Cress, W. D., Triezenberg, S. J. & Greenblatt, J. *Nature* **351**, 588–590 (1991).
- Lin, Y.-S. & Green, M. R. *Cell* **64**, 971–981 (1991).
- Wang, W., Gralla, J. D. & Carey, M. *Genes Dev.* **6**, 1716–1727 (1992).
- Choy, B. & Green, M. R. *Nature* **366**, 531–536 (1993).
- Goodrich, J. A., Hoey, T., Thut, C. J., Admon, A. & Tjian, R. *Cell* **75**, 519–530 (1993).
- Hoey, T. et al. *Cell* **72**, 247–260 (1993).
- Lieberman, P. M. & Berk, A. J. *Genes Dev.* **8**, 995–1006 (1994).
- Workman, J. L., Roeder, R. G. & Kingston, R. E. *EMBO J.* **9**, 1299–1308 (1990).
- Taylor, I. C. A., Workman, J. L., Schuetz, T. J. & Kingston, R. E. *Genes Dev.* **5**, 1285–1298 (1991).

10. Laybourn, P. J. & Kadonaga, J. T. *Science* **254**, 238–245 (1991).
11. Croston, G. E., Laybourn, P. J., Paranjape, S. M. & Kadonaga, J. T. *Genes Dev.* **6**, 2270–2281 (1992).
12. Imbalzano, A. N., Kwon, H., Green, M. R. & Kingston, R. E. *Nature* **370**, 481–485 (1994).
13. Hope, I. A. & Struhl, K. *Cell* **46**, 885–894 (1986).
14. Ma, J. & Ptashne, M. *Cell* **51**, 113–119 (1987).
15. Gill, G. & Ptashne, G. *Cell* **51**, 121–126 (1987).
16. Hope, I. A., Mahadevan, S. & Struhl, K. *Nature* **333**, 635–640 (1988).
17. Kim, Y., Geiger, J. H., Hahn, S. & Sigler, P. B. *Nature* **365**, 512–520 (1993).
18. Kim, J. L., Nikolov, D. B. & Burley, S. K. *Nature* **365**, 520–527 (1993).
19. Horikoshi, M., Yamamoto, T., Ohkuma, Y., Weil, P. A. & Roeder, R. G. *Cell* **61**, 1171–1178 (1990).
20. Arndt, K. M., Ricupero, S. L., Eisenmann, D. M. & Winston, F. *Molec. cell Biol.* **12**, 2372–2382 (1992).

21. Strubin, M. & Struhl, K. *Cell* **68**, 721–730 (1992).
22. Klein, C. & Struhl, K. *Science* **266**, 280–282 (1994).
23. Tzamarias, D. & Struhl, K. *Nature* **369**, 758–761 (1994).
24. Hill, D. E., Hope, I. A., Macke, J. P. & Struhl, K. *Science* **234**, 451–457 (1986).
25. Harbury, P. A. B. & Struhl, K. *Molec. cell Biol.* **9**, 5298–5304 (1989).
26. Myers, A. M., Tzagoloff, A., Kinney, D. M. & Lusty, C. J. *Gene* **45**, 299–310 (1986).
27. Cormack, B. P., Strubin, M., Stargell, L. A. & Struhl, K. *Genes Dev.* **8**, 1335–1343 (1994).
28. Cormack, B. P., Strubin, M., Ponticelli, A. S. & Struhl, K. *Cell* **65**, 341–348 (1991).
29. Brent, R. & Ptashne, M. *Cell* **43**, 729–736 (1985).
30. Collart, M. & Struhl, K. *Genes Dev.* **8**, 525–537 (1994).
31. Chen, J., Ding, M. & Pederson, D. S. *Proc. natn. Acad. Sci. U.S.A.* **91**, 11909–11913 (1994).

ACKNOWLEDGEMENTS. We thank M. Strubin for communicating results before submission of this manuscript, and K. Arndt, M. Horikoshi and D. Tzamarias for providing DNAs. This work was supported by a research grant to K.S. from the National Institutes of Health.

Stimulation of RNA polymerase II transcription initiation by recruitment of TBP *in vivo*

Natacha Klages & Michel Strubin

Department of Genetics and Microbiology, University Medical Centre (CMU), 9 Avenue de Champel, 1211 Geneva 4, Switzerland

EUKARYOTIC transcriptional activators may stimulate RNA polymerase II activity by promoting assembly of preinitiation complexes on promoters through their interactions with one or more components of the basal machinery^{1,2}. On the basis of its central role in initiating transcription-complex formation upon binding to the TATA box, the general transcription factor TFIID, which includes the TATA-binding protein (TBP) and several TBP-associated factors^{3–5}, has been implicated as a target for activators. Consistent with this idea, an increasing number of activators have been reported to bind directly to TBP^{6–9}. To assess the functional importance of these *in vitro* interactions for transcriptional regulation *in vivo*, we made use of a novel strategy in yeast to show that a physical interaction with TBP is sufficient for a sequence-specific DNA-binding protein to increase initiation of transcription by RNA polymerase II. These results imply that binding of TFIID to promoter elements is a limiting step in transcription complex assembly *in vivo*.

The strategy consists of creating an interaction between TBP and RFX1¹⁰, a human sequence-specific DNA-binding protein with no activation potential in yeast (see below), through their fusion to the complementary leucine-zipper motifs that specifically promote dimerization between the c-Myc oncoprotein and its partner Max¹¹ (Fig. 1a). The transcriptional effect of the dimerization-domain-mediated interaction between Myc-TBP and Max-RFX on the activity of a *his3* gene containing a single RFX-binding site upstream of the TATA box is assessed in yeast cells expressing either one or both of these hybrid proteins. Figure 1b and d shows that Max-RFX strongly activates *his3* expression only in cells containing Myc-TBP as a unique source of TATA-binding protein. Activation is strictly dependent on the presence of both an upstream RFX-binding site and a functional TATA element in the promoter of the *his3* gene (Fig. 1c). Surprisingly, whereas Max-RFX stimulates transcription from both +1 and +13 native *his3* initiation sites¹², VP16-RFX selectively increases the +13 transcript (Fig. 1d). Although the molecular basis for this difference is unknown, it suggests that VP16 may activate transcription by a mechanism distinct from the one involved in Max-mediated transactivation.

A critical issue is to exclude the possibility that Myc-TBP contains a fortuitous activation domain, and that it functions as a non-DNA-bound transactivator by being recruited to the *his3* promoter through its association with Max-RFX. To address this point, we investigated whether Myc-TBP behaves as an activator on a *his3* promoter containing the defective TATA element, TGTAAG, in cells expressing the TBP(m3) derivative

that binds TGTAAG with increased affinity¹³. Figure 2a shows that TBP(m3) increases *his3* transcription in cells containing VP16-RFX, but not in cells expressing Max-RFX and Myc-TBP. Hence, Myc-TBP does not function as a conventional activation domain. The finding that Max-RFX and Myc-TBP stimulate transcription from the TGTAAG-containing promoter suggests that the fusion proteins bind DNA in a cooperative manner to allow a productive interaction between Myc-TBP and the TGTAAG promoter element.

To confirm that Myc-TBP is required as a general factor for RNA polymerase II (pol II) dependent *his3* transcription, derivatives specifically compromised in pol II functions were tested for Max-RFX transactivation. The mutants were selected by their ability to support cell growth only in conjunction with a human yeast hybrid TBP (hy17) previously shown to function on pol II promoters, but to be defective for pol I and pol III transcription^{14,15}. Complementation of hy17 requires the TBP mutants to function on pol I and pol III promoters, and their inability to support cell viability on their own would thus be indicative of a pol II specific defect. Two variants bearing mutations within regions known to mediate distinct RNA pol II functions were isolated. One derivative, Myc-TBP(m4), contains the double amino-acid change I194R, L205V which maps on the DNA-binding surface of TBP^{13,16–18}. The second mutant, Myc-TBP(m5), carries a single I230K substitution in helix H2'. Interestingly, a pol II-specific temperature-sensitive mutant has been isolated that contains a single serine substitution at this position¹⁹. Figure 2b shows that, although they are unable to support cell viability on their own, both Myc-TBP mutants complement the growth defect of hy17. This phenotype is entirely consistent with the mutants being specifically defective for pol II transcription. The mutants were then tested, together with the Myc-hy17 chimera, for their ability to respond to Max-RFX. Although Pol I and pol III-defective Myc-hy17 mediates normal Max-RFX transactivation, both mutants fail to do so (Fig. 2b). As the two derivatives bear amino-acid substitutions that are widely separated in the three-dimensional structure^{16–18} and are therefore highly unlikely to both destroy a putative activation domain present in TBP, we conclude that Myc-TBP must bind DNA and function as a basal factor for pol II transcription to respond to Max-RFX. This identifies binding of TFIID on promoter DNA as a limiting step in the assembly of a functional preinitiation complex that may be subject to regulation by activators. By using a similar strategy, it should be possible to assess whether additional step(s) in the assembly pathway following template-binding of TFIID are also rate-limiting *in vivo*. □

Received 19 September 1994; accepted 8 March 1995.

1. Kingston, R. E. & Green, M. R. *Curr. Biol.* **4**, 325–332 (1994).
2. Tian, R. & Maniatis, T. *Cell* **77**, 5–8 (1994).
3. Tanese, N., Pugh, B. F. & Tjian, R. *Genes Dev.* **5**, 2212–2224 (1991).
4. Zhou, Q., Lieberman, P. M., Boyer, T. G. & Berk, A. J. *Genes Dev.* **6**, 1964–1974 (1992).
5. Hernandez, N. *Genes Dev.* **7**, 1291–1308 (1993).
6. Stringer, K. F., Ingles, C. J. & Greenblatt, J. *Nature* **345**, 783–786 (1990).
7. Lee, W. S., Kao, C. C., Bryant, G. O., Liu, X. & Berk, A. J. *Cell* **67**, 365–376 (1991).
8. Lieberman, P. M. & Berk, A. J. *Genes Dev.* **5**, 2441–2454 (1991).
9. Kashanchi, F. et al. *Nature* **367**, 295–299 (1994).
10. Reith, W. et al. *Molec. cell Biol.* **14**, 1230–1244 (1994).

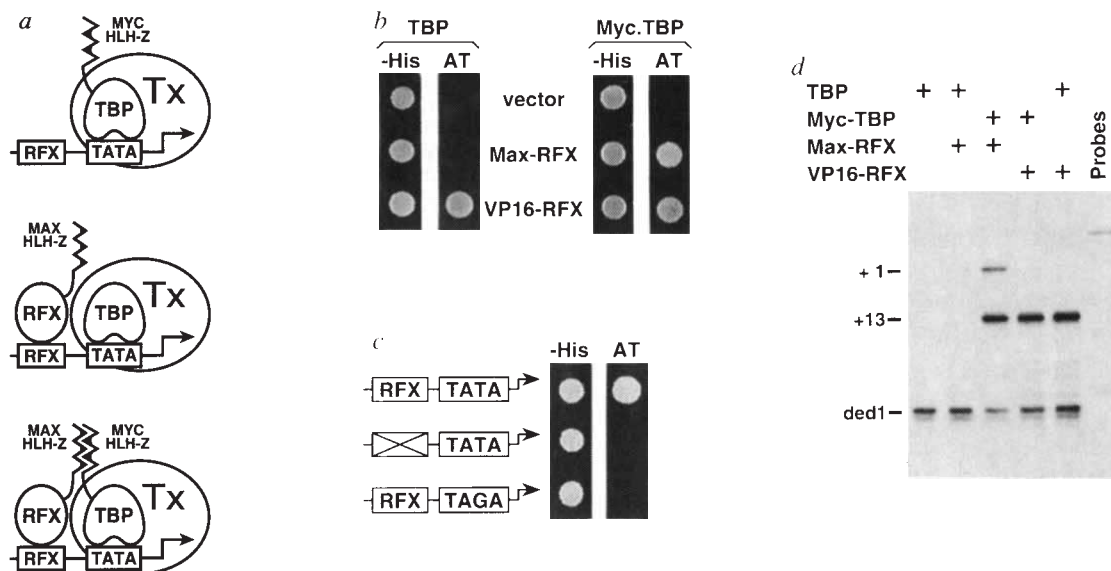
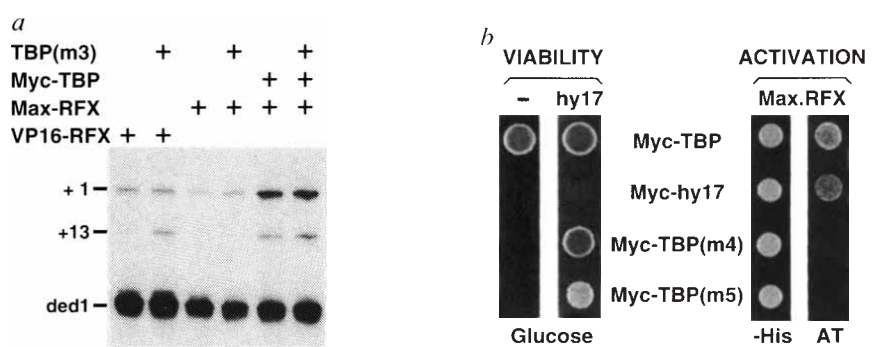


FIG. 1 Recruitment of TBP to a target promoter stimulates RNA polymerase II transcription initiation *in vivo*. **a**, Schematic representation of the strategy. MYC-HLH-Z and MAX-HLH-Z designate the helix-loop-helix-leucine zipper dimerization motifs present in the human c-Myc oncoprotein and its partner Max, respectively¹¹. Tx refers to the transcription machinery. **b**, Strains containing a *his3* allele with an upstream RFX-binding site and expressing the indicated TBP and RFX derivatives from plasmid DNAs were tested for growth on medium lacking histidine (-His) or containing 10 mM aminotriazole (AT). Growth on AT requires induced levels of *his3* expression. **c**, Effect of deletion of the RFX-binding site, or insertion of a point mutation in the *his3* TATA-box (TAGA), on the ability of cells expressing Myc-TBP and Max-RFX to grow on 40 mM AT. **d**, Quantitation of *his3* RNA levels by S1 nuclease analysis. Strains with growth phenotype shown in **b** were analysed for *his3* (+1 and +13 transcripts) and *ded1* RNAs. The level of *ded1* RNA was used as an internal control.

METHODS. All the proteins are expressed from CEN-ARS vectors²⁰ under control of yeast TBP regulatory sequences¹⁴. The Myc dimerization motif derived from pVP-Myc73 (ref. 11) was fused to the N terminus of TBP molecule 4 (ref. 14). Max-RFX contains the EcoRI-SalI fragment from pVP-Max72 (ref. 11) fused to the AgeI site found in the RFX1 coding region¹⁰. VP16-RFX results from the fusion of the VP16 activation domain derived from pMSVP16 D1D3 (ref. 21) to the ATG initiator of RFX. The yeast strains, derived from KY320 (ref. 22), contain *his3* alleles with a single native or mutated RFX-binding site and 50 bp of random DNA inserted at the unique EcoRI site in *his3*-Δ94 carrying the relevant TATA elements²³. Phenotypes were analysed by spotting 10⁴ cells on minimal medium lacking histidine or containing AT. S1 nuclease protection experiments were done as described²⁴, except that the *ded1* radiolabelled oligonucleotide was diluted five-fold before hybridization.

FIG. 2 Myc-TBP functions as a general factor for pol II transcription to mediate Max-RFX transactivation. **a**, Activity of Myc-TBP on a TGATAA-containing *his3* promoter. Strains containing a *his3* allele with the RFX-binding site upstream of the mutated TATA element, TGATAA, and expressing the indicated TBP and RFX derivatives from plasmid DNAs, were analysed for *his3* and *ded1* RNAs. TBP(m3) is the derivative that binds TGATAA with increased affinity¹³. **b**, The indicated Myc-TBP derivatives were tested for their ability to support cell viability either on their own (-), or in conjunction with human-yeast hybrid 17 (hy17)¹⁴, by plating the cells on glucose medium to repress expression of a chromosomal TBP allele under control of the *GAL1-10* promoter²⁵ (left). Their ability to respond to Max-RFX was determined by testing for cell growth on minimal medium containing 40 mM AT. In this case, the medium contains galactose to induce expression of the chromosomally encoded TBP which is required for cell viability (right).

METHODS. The Myc-hy17 chimera was constructed as described for Myc-TBP in Fig. 1. The library of mutant TBP proteins from which Myc-TBP(m4) has been isolated was generated as described¹³, except that



an oligonucleotide containing an equimolar mixture of the four nucleotide precursors at codons 194 and 205 of TBP was used. Myc-TBP(m5) was from a library in which the region of the TBP gene encoding the 14 C-terminal residues was mutagenized at a frequency of 8% per base pair. This was done by cloning a degenerate oligonucleotide into Myc-TBP between the naturally occurring HindIII site and a BamHI site created by PCR downstream of the stop codon.

11. Amati, B., Dalton, S., Brooks, M. W., Littlewood, T. D. & Evan, G. I. L. *Nature* **359**, 423-426 (1992).
12. Struhl, K. *Nucleic Acids Res.* **13**, 8587-8601 (1985).
13. Strubin, M. & Struhl, K. *Cell* **68**, 721-730 (1992).
14. Cormack, B. P., Strubin, M., Ponticelli, A. S. & Struhl, K. *Cell* **65**, 341-348 (1991).
15. Cormack, B. P., Strubin, M., Stargell, L. A. & Struhl, K. *Genes Dev.* **8**, 1335-1343 (1994).
16. Nikolov, D. B. et al. *Nature* **360**, 40-46 (1992).
17. Kim, Y., Geiger, J. H., Hahn, S. & Sigler, P. B. *Nature* **365**, 512-520 (1993).
18. Kim, J. L., Nikolov, D. B. & Burley, S. K. *Nature* **365**, 520-527 (1993).
19. Cormack, B. P. & Struhl, K. *Science* **262**, 244-248 (1993).

20. Sikorski, R. S. & Hieter, P. *Genetics* **122**, 19-27 (1989).
21. Triezenberg, S. J., Kingsbury, R. C. & McKnight, S. L. *Genes Dev.* **2**, 718-729 (1988).
22. Chen, W. & Struhl, K. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2691-2695 (1988).
23. Harbury, P. A. & Struhl, K. *Molec. cell. Biol.* **9**, 5298-5304 (1989).
24. Chen, W., Tabor, S. & Struhl, K. *Cell* **50**, 1047-1055 (1987).
25. Johnston, M. & Davis, R. W. *Molec. cell. Biol.* **4**, 1440-1448 (1984).

ACKNOWLEDGEMENTS. We thank Z. Moqtaderi and K. Struhl for the galactose-inducible TBP gene; B. Amati for plasmids pVP-Myc73 and pVP-Max72; W. Reith for the cDNA encoding RFX1; D. Scherly for oligonucleotide synthesis; and B. Amati, S. Clarkson and W. Reith for critical reading of the manuscript. This work was supported by the Swiss NSF.

CORRECTION

Seismic anisotropy in the mantle beneath an oceanic spreading centre

Donna K. Blackman, John A. Orcutt,
Donald W. Forsyth & J-Michael Kendall

Nature 366, 675–677 (1993)

WE have reduced our estimate of the relative travel-time anomaly that may be associated with anisotropy in the mantle beneath the Mid-Atlantic Ridge at 34° S from 0.6–1.2 s to 0.4–0.6 s. The revised estimate reflects the realization that reverberations between the sea surface and the sea floor introduce serious bias into the long-period pressure data¹. Re-analysis with an

improved filtering scheme suggests that only stations on the east flank show a trend of steady increase in relative travel-time with distance off-axis. It is not clear whether this pattern reflects an anomaly that has an overall increasing trend to the east, independent of age, or whether the low signal-to-noise ratios of all west-flank arrivals have contaminated our results. Pressure-channel arrivals with high signal-to-noise ratio and useful vertical seismometer records are available only from the three east-flank ocean-bottom seismographs.

The lack of an axis-centred anomaly leads us to the conclusion that shear-induced anisotropy associated with mantle flow is consistent with our high-quality east-flank data but it is not required by the complete dataset. If the signal we detect is due to preferred orientation of olivine, our revised estimate suggests a degree of anisotropy (4–8%) more in line with that determined in previous studies (which did not resolve the ridge axis; see, for example, ref. 2) of alignment in the oceanic upper mantle due to plate motion. □

1. Blackman, D. K., Orcutt, J. A. & Forsyth, D. W. *Bull. Seism. Soc. Am.* (in the press).
2. Nishimura, C. E. & Forsyth, D. W. *Geophys. J. Int.* **96**, 203–229 (1989).

KNOW YOUR COPY RIGHTS RESPECT OURS

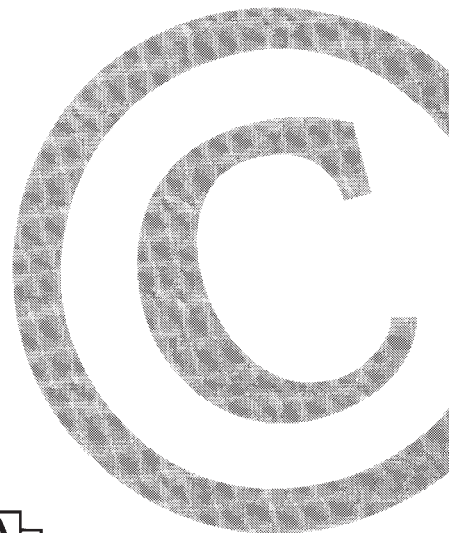
The publication you are reading is protected by copyright law. This means that the publisher could take you and your employer to court and claim heavy legal damages if you make unauthorised photocopies from these pages.

Photocopying copyright material without permission is no different from stealing a magazine from a newsagent, only it doesn't seem like theft.

The Copyright Licensing Agency (CLA) is an organisation which issues licences to bring photocopying within the law. It has designed licensing services to cover all kinds of special needs in business, education, and government.

If you take photocopies from books, magazines and periodicals at work your employer should be licensed with CLA.

Make sure you are protected by a photocopying licence.



The Copyright Licensing Agency Limited
90 Tottenham Court Road, London W1P 0LP
Telephone: 0171 436 5931
Fax: 0171 436 3986

Epochs of nature

Roy Porter

Landscape and Memory. By Simon Schama. HarperCollins/Knopf: 1995. Pp. 652. £30, \$40.

In his typically virtuoso investigation of ideas of landscape, Simon Schama puts history to use to shake up modern beliefs about nature. Endangered species, shrinking icecaps — indeed the shrinking world — have riveted attention on ecological peril. Many feel rightly indignant about endangered species and endangered nature. The Romantic impulse still beats strongly in our breast: the fastnesses of planet Earth — Antarctica, Sun-baked deserts, jungle depths, everything that still resists the intrusive presence of the 'developed world' — continue to exert a mighty appeal. We want 'nature' to be protected against the depredations of 'civilization'.

All of which is very well, but Schama tosses a spanner into the works in the form of a welter of evidence, verbal and visual, demonstrating that the pure and innocent nature we respect and revere is all the stuff of myth. Conservationists worship the English countryside and strive to rescue it from motorways and developers. But there is absolutely nothing 'natural' about that green and pleasant land: our beloved landscape is itself the product of centuries of human settlement, not least the enclosures imposed (ironically enough) by the capitalist agri-business of yesteryear. Marching behind the standard of nature, green activists are often, if truth be known, merely preserving the historical works of man against the present.

And if we can't see this, it is because we remain blinded by the legacy of a Romantic movement that drew a fundamental but false distinction between man and nature and invested the latter with positively religious awe. Victorian poets and philosophers treated nature as sacred and to be honoured for its very wildness, its exclusion of man. Views of that kind made no sense at all before modern times. Feudal society was too busy fighting the forest, protecting itself against feral predators and staving off hunger to be able to indulge in daydreams about the magnificence of the wilderness. The mediaeval forest was a hive of industry, not a retreat of hallowed solitude.

Romanticism blossomed as the luxury of the urban intellectual. And Schama and other critics of Romantic excesses would allege that the same stricture applies to Romanticism's pessimistic heir, contemporary radical ecology. The difference is that whereas Romanticism gloried in the possibility of the poet's rapt communion with creation, 'deep' ecology wants to

debar man from nature altogether.

The roots and ramifications of such crucial issues facing us today are energetically explored in *Landscape and Memory*, a sumptuously illustrated work presenting a dazzling sequence of historical episodes illustrative of man's interplay, in fact and in fantasy, with forests, rivers and mountains. It's a work that resists being reduced to any simple chronological thread and refuses to dole out easy answers. Indeed, as in *Citizens*, his acclaimed account of the French Revolution, Schama revels in the

the Nazis were highly "ecologically conscientious".

Schama's treatment of the Lithuanian forests epitomizes many of his underlying themes. He shows how even (or perhaps especially) in this secular age, our sense of awe when faced with waterfalls or Alpine crags owes much to residual religious impulses. From early on, theologians chose to incorporate elements of the pagan cults of groves and springs; Church Fathers could find God in the desert; St Francis preached to the birds, and healing shrines such as Lourdes attest the lingering appeal of veiled pantheism. In brilliant passages, bedecked with splendid photographs and reproductions of art works, Schama traces such emblems as the 'verdant cross' with greenery creeping around it, symbolically capturing the theological message of the fusing of death and resurrection.

God certainly dwelt in the mountains,



Sunny Morning on the Hudson River by Thomas Cole, 1827. Founder of the Hudson River school of landscape painting, Cole is often viewed as a Romantic artist who invested the American wilderness with religious meaning. From *Thomas Cole: Landscape into History* edited by William H. Truettner and Alan Wallach (Yale, £35).

twists and turns of history.

Opening his grand tour with the Lithuanian woodlands, renowned as some of Europe's last surviving forests, Schama lays bare their shifting fate as the territory was successively fought over by Lithuanians, Poles, Russians and Germans. The forests survive today, not, as might be assumed, as some relic of primaevial times but because over the centuries they have been replanted and sustained as royal hunting grounds — in recent times for the benefit of Nazis such as Hermann Goering, who loved enacting Teutonic fantasies there. Ever one to relish history's wry ironies, Schama, whose Jewish ancestors hailed from the region, notes how

but feelings of popular national destiny were mainly associated with water. Rolling currents have conveyed deep meanings, be it the river of life (latterly the 'stream of consciousness') or the hope of tracing the ultimate source of things — it is no accident that Freud compared psychoanalysis to the quest for the source of the Nile. Fountains comprised emblems of power whereas the smooth still waters of lakes were associated with the soul.

All landscape features are thus saturated with meaning and memory; indeed, they are picturesque, like a picture. Only the fanatical Romantic, preoccupied with his unique encounter with nature, would deny these cultural myths (or what Jung

would call the collective unconscious). Driving round Yosemite National Park in California, it might be tempting to fantasize that here at last is pristine landscape preserved as nature intended. But, Schama notes, the very flora and fauna of the park, with its giant sequoias, is the product of epochs of Native Indian husbandry, involving routine forest burning; and in any case our idea of Yosemite is of a landscape from which the white conservationist, tourist or backpacker has peremptorily excluded its former human denizens.

The face of nature everywhere bears the imprint of man, and nowhere more emblematically than at Mount Rushmore National Monument in South Dakota, with its gargantuan presidential visages carved into the very cliff. Meanwhile, be it in the form of zoos, greenhouses or urban parks, civilization incorporates the spirit of nature. This explains why Schama can feel less gloomy than those radical environmentalists who prophesy the wholesale and final ruin of our planet. For his historical message is that man and nature have been engaged in a perpetual exchange in which man's impact on the landscape creates and recreates environmental aesthetics that perennially assume new meanings and emotional investments. "It is in vain", writes Schama, approvingly citing Henry David Thoreau, "to dream of a wildness distant from ourselves".

Schama is a master storyteller, imaginative and immediate. He is also a master of history in a more academic sense, equally at home with the minutiae of mediaeval theology as when discussing a contemporary artist such as Anselm Kiefer. Not least, he interweaves his historical tapestry with autobiographical threads: the boy living along the River Thames just down from Tilbury, penning a schoolboy history of the Royal Navy; watching his father feasting on whitebait at the Savoy Hotel; and then eventually travelling to the Lithuanian forests to unearth those Jewish roots (implausible though it may sound, his mother's Orthodox forebears turn out to have been loggers).

Like a mature greenwood, Schama's work is dense and variegated; and as with all good fairy stories, the path through the forest is twisting, with impenetrable undergrowth and enticing glades and daring detours. Much is omitted — not least, it must unfortunately be said, the history of science and its impact on Western concepts of the environment. Reading *Landscape and Memory*, the happy wanderer may sometimes feel he cannot see the wood for the trees. But Schama might not be discomfited — he would be ready with a quip, quoting the art historian Aby Warburg: "God lies in the details". □

Roy Porter is at the Wellcome Institute for the History of Medicine, 183 Euston Road, London NW1 2BN, UK.

Monstrous perceptions

Frank Gonzalez-Crussi

The Eye of the Beholder: Deformity and Disability in the Graeco-Roman World.

By Robert Garland. Duckworth: 1995. Pp. 222. £35, \$39. Distributed in the US by Cornell University Press.

At first blush, no human activity seems simpler than seeing. In fact, none is so complex or so precariously perched above the twin-sloped precipice of ambiguity and misperception. But when the object seen is the human body, the risks are multiplied. For a universal notion that we find irresistible is that the body has a representational function; that it is what it is, as



Bronzed beauty — phallic dwarf (early first century AD).

well as 'something else'. It incarnates an interiority, an occult reality, a significance or a 'soul' that projects itself on the visible surface, upon which it leaves, indefectibly, its mysterious mark or imprint. Thus the human body is a cipher, a hieroglyph that must be interpreted; and every age, every society, must provide its particular interpretation. Hence the uniquely apposite title of Robert Garland's new book, *The Eye of the Beholder*, a study of the reactions elicited in the ancient Graeco-Roman mind by the sight of the physically impaired.

Garland is a professor of classical studies at Colgate University, New York. His previous books identified him as an out-

standing scholar with a lively interest in the views held in the ancient Hellenic world on our biological nature, its appalling frailties and inexorable finitude. Garland has now summoned an impressive array of materials to survey the fears, prejudices and superstitions that surrounded bodily deformity, as well as the gallant efforts to transcend the anguish consubstantial to this human tragedy. He draws on vase-painting, sculpture, poetry, drama, ancient medical writings, ethnology, archaeology and mythology. All this is engagingly brought to bear on a learned exploration of the ways in which the ancient Graeco-Roman mind tried to come to terms with the dismal fact of physical incapacitation produced by age or crippling disease.

The picture that emerges is patinated by a certain inevitable callousness, for a society where life was "nasty, brutish and short" — average lifespan of 37 and 44 years respectively for men and women — was not likely to promote warm feelings of genuine concern for the disadvantaged. More commonly, the weak were looked upon with disdain, and contempt was predicated on the unfeeling tenets of a pseudoscientific *physiognomy* that upheld the principle of bodily deformity as an outward expression of inner baseness. Exalted humanists, such as the elegiac poet Theognis of Megara (sixth century BC), bemoaned the fact that the morally depraved are sometimes deceptively handsome, but would not carry this lamentation to the point of challenging belief in the fundamental correspondence between bodily deformity and moral or spiritual wretchedness.

Consequently, the lame, the blind, the ugly and the weak were reviled, turned into scapegoats (*pharmakos*) for the ills that beset the community, then delivered to ritual punishment or execution; released into beggarmdom, sometimes after deliberate maiming, as in Seneca's haunting narrative of children brutally mutilated or blinded in order to increase their earnings as beggars; collected as pets in the household of the rich and powerful, there to carry out menial and humiliating tasks, as did the eunuch whose appointed function was, according to Martial, to steady his master's wavering penis over a chamber pot while he was urinating; displayed publicly as oddities, as were dwarfs and giants during the reign of Augustus; sold as outright merchandise in a 'monster market' or *teraton agora*, mentioned by Plutarch in *Moralia*, and which, according to Garland, may have existed side by side with the market of slaves in many cities of the empire. There, the deformed, coveted as exotic 'objects', could command higher prices than the hale. Martial mentions a buyer who demanded his money back when he found out that a slave he had purchased was not an idiot, as he had

From The Eye of the Beholder

been led to believe.

The physically handicapped were also made into objects of derision, victims of the cutting satire, the pointed, steely jesting and unforgiving irony with which are provided, as with lacerating thorns the rose bush, the comedies of Plautus, Persius and Aristophanes. Garland devotes thoughtful pages to the possible meaning behind the seemingly heartless act of mocking the disabled. Without wishing "to impart any moral judgement into [his] analysis", he points out the complex causality: sadistic impulses, sexual and scatological tendencies, and the desire to exorcise the threat embodied in concrete physical disability, for "to deride the monster is to deprive him, at least temporarily, of his malevolent power". Other chapters deal with the representation of physical deformity in popular or folk art; the groping attempts of the medical profession to correct a few malformations, notably deformities of the spine; the early, rudimentary notions about the cause of congenital malformations; and that most peculiar of collective myths, belief in the existence of 'monstrous races' that populated remote regions of the world.

Garland tells us that we know next to nothing about the fate of infirm or disabled slaves; that the rich epigraphic and literary Greek sources, in contrast to the Roman, hardly mention monstrous births as omens; and that "there are almost no known representations of the severely disabled in the entire canon of classical art", however many vivid depictions there may be in folk art. The eye of *any* beholder, it must be noted, is a zealous gate-keeper. Seeing is not a passive act. We do not retract our eyelids to let the world rush in, as one who opens a sluiceway simply allows the pouring inward of a body of water impelled by purely external forces. No; seeing is an actively exclusionary process, a ceaseless, though perhaps unconscious sifting. Seeing, to function normally, must go hand in hand with blindness. What we fail to see is therefore at least as telling of our intimate self as is the cast and character of our personal vision of the world. Classical Greece chose to lift her eyes in enraptured contemplation of a sublime ideal of bodily perfection, while averting them, perhaps out of a sense of utter powerlessness, from the ubiquitous presence of bodies broken, thwarted and undone.

The Eye of the Beholder is an important, fascinating book that will command the attention of all those interested in the history of ideas. It is an articulate reminder of our unalterable kinship to the historical past. The deformed are no longer regarded as portents or as proper targets of ridicule. But despite our protestations to the contrary, a survey of contemporary history will show that societal and cultural attitudes in the past hundred years have

engendered untold suffering for the physically handicapped. Today, we enter a scientific revolution in the field of genetics still clinging to the belief that the congenitally malformed are, in the words of Garland, "a problem, the curing of which is their final elimination". Two millennia seem not to have decreased one bit the pathos, the gripping power of the well known Sophoclean line, that for some unfortunates "not to be born is, past all prizing, best". To which we continue to rejoin in bewilderment: "Yes, but who is so lucky as to have *that* happen to him? Not one in ten thousand!" Laughter, if devoid of malice, is still a legitimate defence. □

Frank Gonzalez-Crussi is in the Department of Pathology, Children's Memorial Hospital, Chicago, 2300 Children's Plaza, Illinois 60614, USA.

The researcher's dilemma

George J. Annas

Subjected to Science: Human Experimentation in America before the Second World War. By Susan E. Lederer. *Johns Hopkins University Press: 1995. Pp. 192. \$32.95, £27.50.*

THE Nuremberg Code, established by US judges in 1947 at the trial of Nazi doctors following the Second World War, remains the most authoritative statement of the

sary, consent is not a sufficient precondition for lawful human experimentation. Eight of the code's other nine provisions deal with measures researchers are obliged to take before they can even seek a potential subject's consent. Perhaps the most controversial of these is that the "experiment should be designed and based on the results of animal experimentation".

Because of the central importance of the Nuremberg Code, Susan Lederer believes that recent scholarship has tended to ignore the ethics of pre-Second World War experiments and apparently assumed that there were no formal rules before 1947. This seems, for example, to have been the case when the US Advisory Committee on Human Radiation Experiments, charged with establishing the ethics of such experiments conducted by the US government, announced last year that it had discovered a 1953 directive in which the US Secretary of Defense formally adopted the Nuremberg Code as official policy shortly after the election of General Dwight Eisenhower as President. Although this directive, like most military documents, was stamped "top secret", its existence has in fact been well known at least since General Richard Taylor presented it at a public hearing before a US Senate subcommittee in 1975. But neither 1947, 1953 nor 1975 is the critical date in determining when the ethics embodied in the Nuremberg Code should be considered as applicable to all medical researchers. As the code's authors made clear, they were not writing new law or ethics. Rather, they believed they were

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Murderous medicine — Nazi doctors on trial in Nuremberg.

rules of human experimentation despite almost 50 years of pressure from the World Medical Association to replace it with various versions of the Helsinki Declaration. The core of the code is the requirement of the voluntary, competent, informed and understanding consent of the research subject. But although neces-

merely formalizing "certain basic principles [that] must be observed in order to satisfy moral, ethical and legal concepts". So the real issue is whether these human-rights principles were in fact widely held and recognized.

Lederer's research on medical experiments in the United States before the Sec-

and World War supports the conclusion that consent of the subject was a primary consideration in research with humans, at least by 1900. For example, medicine's leading spokesperson on ethics at the time, William Osler, wrote in 1907 that the "full consent" of the research subject was "an ethical requirement". He also argued that although the ultimate test of a new procedure is to try it on humans, that should never be done "before it has been tried on animals". Although the importance that early-twentieth-century researchers attached to consent is well documented in the book, Lederer spends most of her time on her thesis that groups opposed to experiments on animals were the most influential in attacking and limiting experiments on human beings.

At that time, research on both animals and humans was denoted and denounced simply as "vivisection". The most famous essay against the practice was written by George Bernard Shaw in his preface to *The Doctor's Dilemma*. Shaw saw vivisection as simple cruelty that was uncritically justified by the search for knowledge that he believed could be obtained in less barbaric ways. In the United States the groups fighting animal vivisection argued that unless it was prohibited, scientists would quickly move on to human experimentation. Protection of animals was therefore seen not only as a good in itself, but as a way to protect humans from unscrupulous scientists as well. The fact that many human experiments were carried out on orphans and institutionalized mental patients helps to explain the early alliances between animal-protection organizations (humane societies) and child-welfare organizations.

The author's description of experimentation on children and animals before the Second World War is much more complete than her analysis of research on prisoners and members of the military, although those wanting to do their own research on these populations will find places to begin in the book. Her analysis of the legal cases, however, leaves much to be desired. She does discuss the most famous pre-war US case to reach an appeals court, that of *Bonner v. Moran* in which a 15-year-old boy was recruited to try to graft some of his skin to a burned relative by keeping it attached to his circulation (through a "tube of flesh") while the graft was attached to the relative. The author says this case is consistent with the general pre-war US view that physicians experimented on patients at their own peril, being held liable for "adverse results". In fact, the case supports the proposition that reasonable experimentation on mature minors such as Bonner is lawful so long as the informed consent of both the minor and the parents is obtained. That is still the law today.

The antivivisection movement to pro-

tect animals lost much of its momentum during the First World War. Since the Second World War, human experimentation has moved from single researchers working on one or a few subjects to large-scale research projects often funded by government and commercial ventures. The promises of advancing knowledge and conquering disease remain, but the public has become much more supportive of the venture, usually uncritically so. Even in this new atmosphere of public support and enthusiasm, however, the protection of the rights and welfare of human subjects has remained a central concern of society, and a new 'animal rights' movement has arisen that once again condemns gratuitous cruelty to sentient creatures. Those interested in the historical interplay of animal-rights activists and human experimentation will profit from reading this book. Those with a broader interest in human experiments before the Second World War will find it less illuminating. All readers, however, would probably agree with the observation made at the turn of the century by the antivivisectionist Albert Leffingwell: "There is no objection to human experimentation when there is no invasion of human rights". □

George J. Annas is in the Department of Health Law, Boston University Schools of Medicine and Public Health, 80 East Concord Street, Boston, Massachusetts 02118, USA.

No sympathy for the Devil

Mark Pagel

The Lucifer Principle: A Scientific Expedition into the Forces of History. By Howard Bloom. *Atlantic Monthly*: 1995. Pp. 466. \$24.

WHAT little is known of Lucifer's life comes from just a few fragments spoken about him by others, including God. Lucifer left no written work. We do know Lucifer was an angel, but that this career was abruptly terminated. God had to cast him out of heaven for organizing a rebellion among the angels and for stealing light. With what would be irritating consequences for the rest of us, Lucifer landed somewhere below Earth where God had little control over him (Machiavelli would never have made this mistake). This all happened sometime before Adam and Eve. Despite the paucity of information about the man, however, most are familiar with his *oeuvre*.

In Lucifer, writers throughout the ages have found a convenient metaphor for the

evil tendencies that possess human nature. Sociobiology has recently provided a mechanism for this possession: aeons of natural selection have favoured strategies that promote our selfish genetic interests. Although a powerful creative force, natural selection has left us masters at lying, cheating, coveting, stealing, pillaging, raping and murdering. These things are part of the normal routine for the merely infra-human but become evil when expressed in ourselves — presumably because we fear our own worst instincts.

Against this backdrop of the Lucifer-possessed individual, Howard Bloom, a sometime rock-impresario and sometime researcher, constructs the view that humans are nevertheless ineluctably part of a larger social being — the superorganism — for whose ends we will sometimes behave. Superorganisms arise when individuals surrender their own interests to those of a larger group. To Bloom, the cells of a tree, the ants in a colony, the bees in a nest and the humans in a social group act as one, and thereby constitute superorganisms. When it suits the superorganism that we call society, we will wage war on its behalf or even commit suicide when we feel it no longer wants or needs us. Our most terrifying capacities for violence and destruction arise out of our subjugation to the superorganism.

Many of Bloom's arguments will elicit a sense of *déjà vu* to evolutionary biologists, who long ago abandoned the idea that individuals sacrifice their interests to those of a group. One of the great triumphs of evolutionary thinking has been to show how the selfish interests of individuals often coincide with those of the group. Indeed, that may be why the group exists. The peculiar form of genetic inheritance among ants, bees and wasps known as haplo-diploidy means that nonreproductive workers can actually reap higher fitness from rearing their sisters than from reproducing themselves. In other cases, behaviours such as the giving of an alarm call at the approach of a predator, only seem to be for the good of the group. On closer inspection, these alarm calls are shamelessly selfish: they tend either to be given only when relatives are around or to direct attention away from the caller and towards the other (unrelated) fleeing members of the group.

Reproductively inefficient behaviours such as suicide and celibacy may not be performed to advance the group's interests. Bloom uses 'apoptosis' or programmed cell death to understand human suicide. The cells of many multicellular organisms have suicide programmes. When a cell in such an organism no longer receives the message to stay alive, it uncomplainingly activates its death programme and dies. In doing so it probably promotes its genetic interests because the other cells in the organism are clones of it:

the interests of the individual and the group coincide. But it is doubtful whether people commit suicide or give up reproduction to promote a societal superorganism. Human tendencies towards suicide and celibacy could have some ancient basis in promoting one's immediate relatives, if by these actions the relatives are substantially better off. The conditions for this are probably rare and so it should not be surprising that suicide, notwithstanding Jim Jones's obedient followers and the fabled Japanese warriors of the Second World War, is also rare. So is celibacy: America's Shakers — a sect of religious celibates — are on the verge of extinction.

The human capacity for warfare does, on the other hand, deserve some consideration, if only because it is so common and because so many participate in it. Warfare among human tribal societies may have made good genetic sense to individuals if the members of one's own group tended to be closely related. Vast numbers of Crusaders went into battle with little protection beyond the crosses on their chests, and vast numbers died. These men were not necessarily related. But most lived the abject life of a serf, and so promises of salvation and riches may have made crusading seem an even bet. In a contemporary setting, atavistic and utterly repugnant tendencies such as football hooliganism may not so much constitute surrendering of our selfish needs to those of the group, but rather the distasteful expression of them writ large in possibly inappropriate circumstances. But how do we explain contemporary grown-up warfare in which lots of people regularly die? Here it may again pay to examine the influences of poverty, coercion and punishment as much as any role of the superorganism. Few soldiers with other options willingly go to their deaths.

Not many evolutionary biologists would feel qualified to write a book that touches on most things from apoptosis to neural nets to Erasmus Darwin's *Zoonomia*. Fewer still would see a common thread — the Lucifer Principle — linking all these topics to a dark side of human nature. Even fewer would have the courage, as does Bloom, to hold the view that this insight could prevent the decline of the United States in the closing years of the twentieth century. In Nietzsche's *Thus Spake Zarathustra*, the "last man" discovers happiness through rationality. But Nietzsche found this happiness contemptible because it was based on a naive optimism that celebrated rationality (science) as the technique for the mastery of life and the discovery of human values. □

Mark Pagel is in the BBSRC-NERC Ecology and Behaviour Group, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

Sailing the ship of Down

Phillip R. Sloan

Charles Darwin: Voyaging, Volume 1 of a Biography. By Janet Browne. Knopf/Cape: 1995. Pp. 606. \$35, £25.

THE historian of science I. Bernard Cohen once described the "three revelations" of Isaac Newton, as researchers first digested his printed works, then his correspondence and some of his unpublished manuscripts and finally the more recently available archive of unpublished

Peter Brent (1981), Peter Bowler (1990), John Bowlby (1990) and most recently the highly praised study by Adrian Desmond and James Moore (1991), surely alters some basic assumptions about Darwin and his revolution that had prevailed since the Victorian *Life and Letters* and the *Autobiography*. Yet the basic plot of Darwin's life and the cast of characters has remained surprisingly constant. There have been strikingly few truly astounding revelations and no closet manuscripts have come to light to reveal a very different mind at work than one would have expected from the *Origin, Descent* or the rest of the printed corpus. Nor have we yet found evidence of secret lovers or other titillating details of Darwin's

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

The traditional ceremony of 'crossing the line', drawn by Augustus Earle, the *Beagle's* artist. From Captain Robert R. FitzRoy's *Narrative*, 1839.

material. Through these layers of scholarly inquiry, the image of Newton as the archetypal rationalist of the *Principia* was replaced by the Newton of the prophecies of the Book of Daniel and the "Queries" to the *Opticks*, and ultimately by the esoteric alchemist in search of the "Greene Lyon".

The potential for a similar transformation in our understanding of Charles Darwin has existed since 1959 when the full archive of his manuscripts and letters became available. The heroic efforts of the notebooks and correspondence projects have drawn together a remarkable body of information. Darwin promises to become the most completely documented scientist in recorded history. But who is the Darwin of this third revelation?

The run of biographies and partial biographies that has accompanied this new archival scholarship, beginning with Gertrude Himmelfarb's *Darwin and the Darwinian Revolution* of 1959 through to the studies of Edward Manier (1978),

private life.

The mysteries and gaps in our understanding of his scientific thought still need clear explanations. When and why did he first think of transformism? What were the reasons for his refusing to publish details of his theory in 1844? Why were barnacles worth so much of his time and effort? Something more modest has instead been attained: the assemblage of a vast amount of material for describing in remarkable detail the intimate picture of the life and times of a largely conventional Victorian intellectual living through a period of dramatic social and scientific change.

Yet we still do not really have a true 'scientific' biography of the genre represented by the classic works of Richard Westfall on Newton, Thomas Hankins on Jean D'Alembert and William Rowan Hamilton, and Pearce Williams on Faraday, in which the science is brought to life by the use of personal biography. The problem may have been made more diffi-

cult by the extensive work of the Darwin scholars themselves. The mass of archival material, correspondence and commentary now accumulated around the man has become so immense that it is almost impossible to condense.

Janet Browne's splendid book is the first volume of a patient analysis of Darwin's entire life in the light of all these sources. Her qualifications as a trained biologist, historian of science and skilled editor of the correspondence project put her in an ideal position to do this challenging work. To those aware of contemporary Darwin studies, the events, characters, work and basic plot of Darwin's life up to 1856 remain generally unaltered in her study. The novelty is in its emphasis on issues and individuals ignored by others and on the scientific dimensions of Darwin's life. Browne's account is strictly chronological, beginning with the young gentleman at the Mount in Shrewsbury in the world of Jane Austen. We follow him from his chemistry experiments with his elder brother Erasmus to his medical (and natural history) studies at Edinburgh, Cambridge, the *Beagle* voyage, London and finally to Down House where the book ends with Darwin's return in May 1856 to the "species question".

Those looking for history with an agenda will be disappointed. If Browne makes Darwin into a product of his times, a man whose "theory of natural selection could only have emerged out of the competitive context of Victorian England", she does not try to reduce the content of his scientific work to the interplay of Victorian social forces and class conflicts. And she makes little effort to psychoanalyse her subject. Even Darwin's debilitating illness, which has tantalized both parasitologists and professional psychoanalysts, lurks in the background except when it interrupts his scientific work, as with his three-month stay in 1849 at Malvern where he was initiated into the rigours of Dr James Manby Gully's famous water cure.

Browne sensitively reveals the person behind the correspondence and notebooks. She is scrupulously faithful to the texts and the documents in constructing her narrative. Where evidence is slight, such as in the case of Darwin's possible liaison with Fanny Owen of the "Housemaid and Postillon" letters, she simply admits that there is little to conclude, and on the whole allows the social proprieties of the Regency world to decide the issues. Her analyses are exact, plausible and eminently judicious.

Browne's most important 'revelation' is perhaps the important role finally given to Charles's elusive elder brother Erasmus, his first scientific mentor. Erasmus's preparation of the way for Charles's initiation into the Cambridge network has previously received little attention. Their

close relationship in the years between Charles's return from the *Beagle* voyage and his marriage to Emma Wedgwood shines through in the correspondence, but it has needed expanding and deepening. Browne has fleshed out Erasmus's character in fine detail, portraying his circle of associates, his scientific and other intellectual interests and his curiously flaccid personality.

Browne's earlier book on the history of biogeography (*The Secular Ark*, 1983) and her training as a biologist equip her well to examine Darwin's years aboard the *Beagle*, the subject of eight central chapters. Her blending of the geological and biological findings is deft and informative. Darwin's anthropological observations are also drawn together into a coherent and engaging narrative, setting the stage for the incorporation of 'man's place in nature' into his transformist theory.

The book is aimed at a general audience. Darwin specialists may wish for more detail about the crucial scientific inquiries during the *Beagle* years. And although the general framework and balance of the discussion is excellent, it would have been strengthened by more judicious use of quotations and illustrations from the unpublished zoology and geology manuscripts and the field notebooks and manuscripts. The *Beagle* years are typically read in historical reverse in most studies, as if the final cause of the voyage were the creation of the transmutation theory. We need to grasp the problems and issues that were then actually at issue for the young Darwin, problems and issues that may have had little direct bearing on transformism but which might explain many other aspects of his later thought. Browne certainly does not engage in such retrospective readings, and her grasp of the historical documents and literature is superb. But she could have introduced the novice a little more to the scientific, as well as the personal, details.

Voyaging is above all a wonderful read. Browne ends with the metaphor of Down House in the 1840s and 1850s as a kind of ship, almost a re-creation of the *Beagle*, self-contained, with specimens and bottles in order, and children and servants replacing Wickham and Covington as aids in Darwin's research. Over it all stands the Victorian father as a disciplined and able captain, continuing in his study the work begun on the large table of the *Beagle* wardroom, with all domestic functions subordinated to a scientific mission. The outcome would be one of the more remarkable products of human genius. I look forward to the sequel. □

Phillip R. Sloan is director of the Program in History and Philosophy of Science, University of Notre Dame, Notre Dame, Indiana 46556, USA.

Paperback history

■ **The Path to the Double Helix: The Discovery of DNA** by Robert Olby, with a foreword by Francis Crick. Dover, \$13.95, £13.95.

The Eighth Day of Creation: Makers of the Revolution in Biology by Horace Freeland Judson. Penguin, £12.

First published in 1974 and 1979 respectively, these classic books provide in-depth accounts of the events leading up to the discovery in 1953 by Watson and Crick of the molecular structure of DNA.

Judson, a renowned popularizer of science, is the better writer; and, unlike Olby, he goes on to look at the dramatic development of the basic concepts of molecular biology in the 1950s.

In his new postscript, however, Olby, a noted historian of science, describes what he perceives to be the different approaches of the two versions: "For him [Judson] Kuhn's 'celebrated fashionable discourses on scientific revolutions' are not well informed on the practices of scientists. Judson prefers to treat scientific development in terms of the 'correspondence principle', in which the inherent conservatism of science is stressed. A new theory has to explain all that was explained by the old one. In this way the cumulative nature of scientific knowledge is preserved. The problem with this view is that it fails to accommodate rejected knowledge. Such knowledge has to be treated as anomalous, ambiguous, mistaken. . . Nor need we exclude tentatively held and vague theories simply because they were advanced with less conviction than were other more successful theories."

■ **The Physicists: The History of a Scientific Community in Modern America** by Daniel J. Kevles. Harvard University Press, \$17.95, £14.25. On its original publication in 1971, this book was heralded as an indispensable introduction to the institutional history of US science. Kevles, a historian of science, traces the social, cultural and political forces behind the growth of physics from the post-Civil War years to the present. The new edition contains a 34-page account of the rise and fall of the Superconducting Super Collider.

Indecent exposures

June Goodfield

Marie Curie: A Life. By Susan Quinn. Simon and Schuster: 1995. Pp. 509. \$30.

WHEN I was a schoolgirl during the Second World War, the only available role model for an aspiring woman scientist was Marie Curie. Judging from the citations in dictionaries and encyclopaedias of scientists, even as late as the 1970s, it would indeed have seemed that she was the only woman scientist. Awe-inspiring, brave, poor, she had a hard life. An emigrée from her native Poland which she adored, she devoted her efforts to physics and mathematics and to her colleague and husband, Pierre. If anything, her devotion and single-mindedness became more intense after the tragedy of his death. She acquired two Nobel prizes from Sweden and a few critical grams of radium from admirers in the United States. She continued to pay a high price for her selflessness and single-minded dedication, for even her final illness and death were provoked by exposure to radium, the new element she had discovered.

Our only source of popular reading about her was the biography by her daughter, Eve, published in 1937, three years after Marie died. The speed with which it was written was astonishing and quite deliberate. As Eve Curie Labouisse confessed to Susan Quinn 51 years later, she wrote quickly because she was "afraid that someone else would do it first and not get it right". This was not the first time — nor will it be the last — that a biography of a famous person has been assembled in a great hurry by relatives or devoted friends in order 'to get it right'. Sometimes they even do it themselves, as Margaret Mead once told me she had done with her autobiography, *Blackberry Winter*. Moreover, as Quinn points out, the portrait of Marie Curie that came down to us was not only a tremendous idealization but also served as a defence by Eve Curie of her mother.

Such a defence was necessary because Marie Curie's affair with a fellow scientist, Paul Langevin — married and the father of four — led to her public exposure and dreadful vilification. Although dismissed by Eve Curie in a few paragraphs, this episode was not only a major trauma for Marie Curie but also reflected many things: the passionate woman whose objectivity could collapse — as most people's can — when the heart and emotions are closely involved; the crucial

attitudes of fellow scientists; and the social milieu in which she worked and lived.

Fifty years ago, myth and idealization, nobility and dedicated objectivity were the hallmarks of science. The image of the noble scientist — mostly men of course, plus Marie Curie — was not only endorsed by scientists themselves but also

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Le Petit Journal marks Marie Curie's visit to the United States in 1921.

often positively promoted. This noble edifice began to crumble almost in direct proportion to the speed with which the enterprise of science, and the number of its practitioners, began to increase. In the United States this was due to the establishment of the National Science Foundation and the National Institutes of Health and to the administration of President John F. Kennedy. The ultimate financial commitment was a pork barrel of such wealth that Jerome Wiesner was once moved to observe that he had never seen one of such dimensions, the real difference this time being that "the pigs were running the show". As the money available increased, so did the competition for it. As the competition intensified, so the ethics, the image and indeed the actions of some of those very 'noble' — as well as Nobel — scientists were eroded. And with the publication of *The Double Helix*, that

deliciously indiscreet, sometimes very unfair but frank account of the discovery of the structure of DNA, Jim Watson gave a new dimension to the image of the scientist, with the magisterial imprimatur of one who had gained the most coveted success of all.

But if the ambience and image of science have changed, so too have other elements of the profession. Science now operates not only with many more women scientists than ever before, but also in an atmosphere that is strongly feminist — at least in the United States. And this provides the main motive for Quinn's biography. For as she writes: "my reasons for undertaking a biography of Marie Curie

were as much of our time as Eve Curie's were of hers. . . I have looked for evidence that Marie Curie was not just a singular, exceptional woman (though she was indeed that) but also a woman who experienced the same difficulties as other women with strong opinions and ambitions."

Then, as now, there were many barriers to women's advancement in the profession. Then, as now, there were defeats and humiliations. In Marie Curie's case these came thick and fast at the hands of the French Academy of Sciences, the bourgeoisie and an "outrageous right wing press".

Quinn has produced a magisterial piece of work. Admittedly she has had access to sources not available until 1990 — notably Marie Curie's own journal. She has also delved deeply into family documents and into the archives in Warsaw, Paris and Sweden. The result is wonderful in its depth and detail, so much so that it will be years, if ever, before this account is displaced. Only new documents and new evidence could do that.

Even then the book will probably never be superseded, only complemented.

Not only is this a biography that tells you everything that you need to know; it also succeeds, quite brilliantly, in a conceptually new way. By seeking to understand Marie Curie in terms of attitudes and questions relevant and prevalent at the end of the twentieth century, the author has gained a profound understanding of a woman scientist at the turn of the nineteenth century. This goes to show that in biography, just as much as in science, asking the right question is the beginning of wisdom. The balance too is right, beautifully judged between the science, the personality and Marie Curie's own belief that the most important part of her work was not the "much touted cure for cancer" but her "critical insight that radioactivity was an 'atomic property' of her newly discovered elements", the most

Mary Evans

important precursor to our modern understanding of atomic structure.

She strove throughout, and especially at the end, to keep a firm distinction between her personal and scientific life. But in this she failed, and a woman of greater understanding of the world would have realized that she was bound to fail.

She was awarded the Nobel prize for chemistry at the height of the open scandal over her affair with Langevin. A member of the Swedish Academy wrote to her indicating that she would not be welcome in Sweden and should refuse the prize until she had cleared her name. But, she replied, the prize was given for her discovery of polonium and radium, and nothing else. Was she right to insist — is any scientist right to insist — that there is “no connection between scientific work and private life”? Given the facts, and that she had written an incredibly indiscreet letter to her lover, with detailed recommendations as to how he could withhold sexual favours from his wife and thus make a break inevitable, there was probably no way that Marie Curie could be treated fairly by contemporary French society.

There is something very Janus-faced about the situation and, for an English reviewer, the furore and scandal that surrounded Marie Curie's affair in France is very hard to understand. For I am writing at a time when the president of France not only has an illegitimate daughter but is also applauded for his paternal devotion. This contrasts sharply with the speed with which a number of English politicians have resigned for extramarital affairs. Some have an illegitimate daughter as well, in one case two of them. As Quinn points out, in France certainly, bourgeois men could keep — and still can keep — a mistress so long as she stayed in the background, as did President François Mitterrand's. That enforcer of male privilege, the Napoleonic Code, was indulgent towards the husband. But Marie Curie was in no way anonymous; she could not fade into the background. She had a career, an independent income and ambitions, and was therefore completely vulnerable to public exposure. Her letters to her lover were stolen and published, letters fuelled by passion as fiery as the passion that, as Einstein pointed out, she demonstrated at scientific conferences.

Marie Curie's impotence in the face of the ‘outrageous press’ was total, as apparently is that of adulterous British politicians and all the rest. A touch of farce attended the affair when Langevin challenged Gustav Téry to a duel for insulting him in an article that accompanied the publication of the letters. Neither of them ever intended to fire the pistols, nor did they, but the duel was the talk of Paris and of Sweden too. It was shortly after this event that Marie Curie was recommended not to go to

Stockholm to receive the Nobel prize.

The episode affected her profoundly, of course, both personally and scientifically. It ruined her chances both of becoming a member of the French Academy of Sciences and of starting a new life with her lover. He was reconciled with his wife and took another mistress (an anonymous secretary) while Marie Curie was left to go on alone.

Whether about this scandal, or the details of her scientific discoveries, or her theories or her childhood, the material in this volume is impeccably researched and splendidly presented. This is not a book that one devours at a sitting as one did *The Double Helix*. It is far too profound and thought-provoking for that. Yet I also found it, and its subject, devoid of humour, as is Eve Curie's biography. True, there are lighter touches. As is usual with all good discoveries, quacks and opportunists raced to exploit them, trading on “the assumption that water's radioactivity had health giving powers and the Curies' good name”. This lasted for a long time. I remember, again from my earlier years, drinking mineral water in France and seeing the label on the bottle list the quantities of magnesium, zinc and all the other goodies therein, and the radioactivity too. How quickly radioactivity disappeared from the list. So it is not surprising that there was a “Curie hair tonic” that was claimed to stop the loss of hair as well as restoring its colour and a “*creme activa*” that held out the promise of eternal youth with the statement that “Madame Curie... promises miracles”.

All this was in deplorably bad taste of course. But the puzzle of Marie Curie persists. While she may have had her lighter moments, irreverent humour was never one of her strong characteristics; and perhaps it doesn't matter for scientists, other than that those without a sense of humour will have a hard time. If, on looking back, I realize I didn't have what it takes to be a scientific nun, it's quite clear from reading this book that Madame Curie didn't have it either. As a role model, the injection of Eve Curie's book provoked an immunity to dedicated laboratory work in me, although it had the opposite effect on my admirable sister.

So how, in 1995, would the Madame Curie portrayed in this biography shape up as a role model for today's aspiring women scientists? Although I don't really know, I suspect the young women of today will find more inspiration in the deliciously eccentric and formidable Barbara McClintock, who brilliantly and comprehensively gave those Young Turks of molecular biology their come-uppance. For we are all creatures of our times. □

June Goodfield is at International Health and Biomedicine, The Manor House, Alfriston, East Sussex BN26 5SY, UK.

My families and other animals

W. C. McGrew

Reflections of Eden: My Life with the Orangutans of Borneo. By Biruté M. F. Galdikas. Little, Brown/Gollancz: 1995. Pp. 408. \$24.95, £16.99.

Now comes the long-awaited volume to complete primatology's most famous ‘trilogy’: Biruté Galdikas's account of her long-term field study of the orangutans of Kalimantan in Indonesia. She joins Jane Goodall (*In the Shadow of Man*, 1971; *Through a Window*, 1990) and the late Dian Fossey (*Gorillas in the Mist*, 1983) in



From *Reflections of Eden*

Galdikas: self-reflection of reflected self?

giving a personal report of behavioural research on great apes in nature. Goodall has studied the eastern chimpanzees of the Gombe National Park in Tanzania since 1960; Fossey looked at the mountain gorillas of the Virunga volcanoes in Rwanda from 1967 until her death in 1985; and Galdikas has focused on the Bornean orangutans of Tanjung Puting National Park since 1971.

What makes the project part of a trilogy is the common source: the Anglo-Kenyan palaeoanthropologist Louis

Leakey, who assigned to each of these women a species of great ape. (Actually, we are told, Leakey gave Galdikas a choice: she could have had the pygmy chimpanzee or bonobo, but she stuck to the Asian ape.) Leakey's influence as the 'spiritual father' of these scientific siblings is immense and pervasive: much of the final, summing-up chapter is a synthesis of Galdikas's feelings about her 'sisters'. As the youngest offspring, Galdikas seems to be the most keen to uphold Leakey's memory, perhaps because he died before she achieved success.

Unlike the books by her fellow tri-mates, *Reflections of Eden* is highly ego-centric, sometimes frustratingly so. Of the 64 photographs, 37 feature Galdikas, whereas only 36 portray orangutans. (Oddly enough, there is no photograph of the forest itself, despite the book's title.) No references are given should the reader want to pursue Galdikas's scientific writing or the alternative views of others. There is no index, and most of the 22 chapters have enigmatic one-word titles, usually proper names. Chronologically, the book concentrates on the first four years of the project — as late as Chapter 16, we have reached only 1975.

The substance of the book is a memoir of what it is like to study wild orangutans and to live with ex-captive orangutans who have been confiscated and then released back into the wild. Following Goodall's and Fossey's lead, Galdikas concentrates on individual apes, about half from each category. Here Galdikas is at her best. She is a master storyteller: vivid, evocative, moving. In delving into the hearts and minds of her subjects, she is intuitively persuasive. (Interestingly, unlike most field primatologists, her undergraduate major was in psychology.)

Underlying the account is a spiritual theme that goes beyond the recurring imagery of Eden. God features prominently, most startlingly as a super-hominoid player of cat's cradle. A walk into the rainforest is likened to a walk into the mind of God. Angels keep popping up, either as wealthy donors in Los Angeles or as self-descriptors for the trio, with, for example, Fossey's anti-poaching efforts making her an avenging angel. Galdikas tells us that her revelatory calling came in the form of a crystal-clear chime during a lecture at the University of California, Los Angeles. So it is not surprising that the book's final sentence reads: "We are allowed to see the eyes of God [when we look into the eyes of an orangutan]."

Happily, the scientific findings after decades of research are notable. Orangutans really are loners: Galdikas has never seen two adult females groom one another — and they are reckoned to be the more sociable of the two sexes! Males disperse from their natal ranges, whereas females stay at home, making the Asian

ape different from its African cousins. The typical birth interval is eight years, the longest of any species of primate. Despite the remarkable ingenuity and imitatorial skills of released captive orangutans, such as their attempts to kindle cooking fires, their wild counterparts show no subsistence technology whatsoever. They regularly mate face to face and engage in forcible copulation, which Galdikas likens to "date-rape".

Unfortunately, knowledge from her impressive research is rarely integrated with that of others. Of her contemporaries, only John MacKinnon is credited. Pioneers such as Peter Rodman and David Horr are mentioned only to be dismissed, and decades of research by Dutch workers such as Hermann Rijksen on the other subspecies of orangutan in Sumatra are simply ignored. Japanese researchers (Akira Suzuki, for example) and Indonesian researchers (such as Jito Sugardjito) suffer a similar fate. Galdikas may well have spent more time near wild orangutans than all other primatologists put together, but science is not a solitary activity, even when its subjects are.

The strengths and weaknesses of the book ultimately boil down to one basic difference between Galdikas and her counterparts, Fossey and Goodall. Unlike them,

Galdikas undertook to do both research and conservation from the beginning. Within a week of arriving in Borneo, she instigated the rescue of a pet orangutan and so began a commitment to individual welfare that persists to this day. Her field site, Camp Leakey, is thick with ex-captives, from panhandlers to irregulars. Rehabilitation of infant great apes is a round-the-clock job, exhausting in all ways. To have done it in parallel with full-time study of wild primates is unprecedented, like holding down two jobs, each of which could be all-consuming. This double life has taken a toll along with the rewards: at no point does Galdikas seriously address the key issue of the possible impact of her immigrants on the lives of the wild resident apes. More than a hundred incomers have been released into what was apparently an ecosystem already at carrying capacity, yet she fails to consider the consequences of enhanced competition. At the end of the book, one wishes it were two, one about the natural lives of our close relations and another about the challenges and choices of repairing our human mistreatment of those same cousins. □

W. C. McGrew is in the Departments of Anthropology and Zoology, Miami University, Oxford, Ohio 45056, USA.

A river runs through it

Christopher Wills

River Out of Eden: A Darwinian View of Life. By Richard Dawkins. *Basic Books/Weidenfeld and Nicolson*. Pp. 166. \$20, £9.99.

THIS short book is the latest in the Science Masters series, a set of brief explorations of their fields of expertise by some of today's most distinguished science writers. Richard Dawkins treats the subject of evolution in his usual limpid style. The book breaks no new ground but, as usual, it abounds with metaphors that make things brilliantly clear. As someone in the metaphor business myself, I must admit that nobody can turn a metaphor better than Dawkins. The central metaphor here is that evolution is like a river, flowing smoothly (more about this later) and made up not of water but of bits of digital information. We are rapidly moving into a digital world, and Dawkins, no lag-gard, points out that it is lucky that our genetic material is digital rather than analogue.

It is not impossible to imagine an analogue mechanism for passing genetic information from one generation to the next. Suppose a gene were to consist of a protein molecule of a specific shape,

which serves as a template for the construction of another protein containing a negative image of it. This in turn would serve as a template for the construction of a molecule of roughly the original shape, and then the process would be repeated. It would not take many generations before the genetic message would fade into indecipherability. And, although Dawkins does not make the point, that is how geneticists tended to think about gene replication before the age of Watson and Crick.

I immediately tried out Dawkins's picture of a flow of digital information on students in my class on molecular evolution. Most of them, it turned out, knew all about digital information, but few had come across the concept of analogue information. (Those who want to use his image had better hurry, before the memory of analogue watches and analogue records fades from the collective consciousness.) Dawkins uses the metaphor of a digital river to link together lucid discussions of several fairly disparate subjects: the provenance of the mitochondrial Eve, the way in which complex organs and behaviours might have evolved and the problem of good and evil. His view of the last will be, for many potential readers, a bone-chilling one:

"The universe we observe has precisely the properties we should expect if there is, at bottom, no design, no purpose, no evil and no good, nothing but blind pitiless indifference. . . . DNA neither knows nor cares. DNA just is. And we dance to its music."

This is a bit too remote. It is the dance, of course, that is important to us. It produces events that, if not good or evil, are certainly excellent imitations. There is also a certain remoteness about Dawkins's metaphor of a river. His image of a digital river is, it would seem, coloured by his experience of the well-tamed genteel rivers of the English countryside. There are no rapids to roil the flow, no chasms and sudden changes of direction. Yet it is becoming increasingly apparent that dramatic events can happen to the digital river, and it would have been exciting if he had explored some of them.

He goes too far in rejecting the claims of the proponents of punctuated equilibrium.

Emerging new science

Ian Stewart

Thinking in Complexity: The Complex Dynamics of Matter, Mind, and Mankind.

By Klaus Mainzer. Springer: 1994. Pp. 329. DM58, \$58, £23.

SYNERGETICS, non-equilibrium thermodynamics, catastrophe, symmetry-breaking, chaos, fractals, self-organized criticality, antichaos, complexity. . . . For as long as I can remember, small bands of scientists led by maverick gurus have chopped away at their own little corners of a big problem — the occurrence of complex structures in what ought to be a simple Universe. One of their common themes has been the emergence of 'collective phenomena', the insistence that the whole is not so much greater than the sum of its parts as different from it.

The deeper theme, however, is non-linearity. Huge areas of traditional science are based on linear thinking, often without explicitly recognizing it. Classical population genetics, for example, is linear: it models a population as a homogeneously mixed pool of genes and studies only the proportions of particular genes. It is a 'mean field' theory that bases its entire conceptual armoury on a convenient fiction. Traditional mathematical economics is also relentlessly linear in viewpoint because of its emphasis on equilibrium behaviour and a one-dimensional additive measure of value. An instructive example of the sheer vapidness of this way of thinking is a recent comparison of the cost of global warming with the cost of preventing it, carried out by the Intergovernmental Panel on

Even though the idea as originally proposed had no mechanism attached to it, more and more mechanisms are being found that can speed up and slow down evolution. Walter Gehring and his collaborators can make compound eyes appear all over the body of a fruitfly, like a rash. This tells us something about the capability of a fly's genome to do startling things, a capability that is already present and that Gehring can tap into.

The book, although an excellent introduction to many important evolutionary ideas, slightly suffers because of Dawkins's reluctance to discuss some of these fascinating recent breakthroughs. Had he done so, the river that emerged might not have been so limpid but it could have provided a wilder ride. □

Christopher Wills is in the Department of Biology, University of California, San Diego, La Jolla, California 92093-0116, USA.

Climate Change (IPCC). Among other remarkable assumptions, the study valued the life of a European at ten times that of the life of an Asian, concluding — although not in quite these words — that it makes economic sense to let Bangladesh sink beneath the waves to avoid minor discomfort in England.

There are many places where linear thinking works well, otherwise it would not have survived, but a cruel irony has led to its widespread adoption in many areas where it does not. Scientists reasonably work with what tools they have to build the biggest edifices they can manage. Linear mathematics is easy, whereas nonlinear mathematics was — at least until recently — impossible. So scientists erected obelisk monuments to linear thinking, blissfully unaware that nature is largely nonlinear.

The times, they are a-changing. Non-linear modelling is one of the biggest growth areas in the whole of science. To those of linear upbringing it may seem undisciplined — and it is, but only in the sense of its 'not belonging to any particular discipline'. It is an irreducibly interdisciplinary way of thinking and it slices the cake of science in totally new directions.

Klaus Mainzer argues the case in favour of nonlinear thinking across the scientific board, from quantum mechanics to human society. He is unusually strong on history, taking care to place each argument in its proper historical context. This is an effective technique, driving home the fact that nonlinear methodology has its roots in ancient debates about matter, life and the mind.

Our deepest theory of matter, quantum mechanics, is linear — indeed it is probably the most spectacularly successful linear theory we have ever had — but the process that turns a quantum system into a classical measurement is manifestly not. So something nonlinear is going on, which cannot be captured by Copenhagen-style special pleading about collapsing wave functions.

Life poses irreducible problems for linear thinking. If one crushed all living creatures together with a huge mortar and pestle, the resulting chemical mix would show few of the characteristics of life (although it would mimic a linear 'gene pool' superbly). Life is a non-linear process of increasing complexity, explicable in terms of dissipative self-organization.

With regard to the mind-brain problem, Mainzer has little time for either Descartes or Penrose. His approach is best summed up by a direct quotation: "The emergence of mental states. . . is explained by the evolution of (macroscopic) order parameters of cerebral assemblies which are caused by nonlinear (microscopic) interactions of neural cells in learning strategies far from thermal equilibrium". This thesis is developed in greater detail in the ensuing discussion of artificial intelligence and the self-organizing and learning abilities of neural nets.

Finally we come to the problem of political, social and economic order in human society. Linear thinking views these things in mean-field terms, for example in concepts such as 'the inflation rate' and 'the unemployment rate'. Note those 'the's'. In a real economy, individuals suffer their own variable inflation rate and their own even more variable unemployment rate. The macro-variables are emergent features of a complex system of millions of interacting agents, each with its own micro-variables. In a classical linear economy, stock markets never crash.

All very well, but what should we actually do? The epilogue has some suggestions. Mainly they are to recognize the interdependence of systems that we usually try to keep separate (such as economics and the environment) and to accept that complex systems develop emergent properties — in short, to stop thinking linearly. Nonlinearity is not a universal answer, but it is often a better way of thinking about the problem. It is certainly better than the linear thinking that led the IPCC to conclude — no doubt without realizing it — that it makes sense to let nine Asians die in order to keep one European alive. □

Ian Stewart is at the Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK.

Technical boundless optimism

David E. H. Jones

Nano! Remaking the World Atom by Atom. By Ed Regis. Little, Brown/Bantam. Pp. 307. \$23.95, £16.99.

THE boundless-optimism school of technical forecasting covers a wide spectrum. The more realistic end features such marvels as nuclear electricity too cheap to meter (remember that?) and computers that understand spoken English. At the manic end we find space colonies located at Lagrangian points, the terraforming of neighbouring planets, and robots in convincing human form. A promising newcomer has recently appeared somewhere in the middle of this spectrum. It is called nanotechnology.

Nanotechnology must be carefully distinguished from microtechnology. The latter, which could operate down to the micrometre scale, is simply the manufacture and use of extremely small mechanisms — microscopic lathes, motors and so on. Nobody has got very far with it, and it might or might not be much use, but the theory seems unexceptionable. As early as 1959 Richard Feynman pointed out that a lathe could be used to make the parts for a smaller lathe; this could make the parts for one smaller still, and the process could continue downwards to no very obvious limit. Modern photofabrication can get quite a way in one step, by exploiting the techniques developed for integrated-circuit manufacture. Micrometre-sized microphones and force-transducers have already been etched out in this way.

Nanotechnology takes this idea to its ultimate extreme — the direct assembly of components and structures atom by atom. Its prophets assert that anything whatever could — and once the trick has been perfected, will — be made by atomic assembly. The nanotechnological John the Baptist is Eric Drexler of the Massachusetts Institute of Technology. He has been crying in the wilderness since at least 1980 and has acquired a group of enthusiastic disciples. Ed Regis is one of them, and *Nano!* is his account of the revelation so far.

Nano! is a dramatic work. It is a breathless, exhaustingly complete, desperately confusing, utterly uncritical, thought-by-thought, meeting-by-meeting, paper-by-paper account of Drexler, his life, times and growing missionary conviction that nanotechnology will lead us to the promised land. As a story, a personal biography of an imaginative and committed individual, it has its own interest and

human appeal. As an account of a clique of enthusiasts and their vision of the future, it sheds an intriguing light on the psychology of technical boundless optimism.

The vision, roughly, is this. In the nanotechnological world, everything will be made by tiny programmable 'assemblers', about the size of bacteria, whose exact specification cannot yet be given. Swimming around in a sort of water-bath,

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Small comparison — unlike nanotechnology, the theory of microtechnology seems unexceptionable. Shown here are micromechanic components next to a fly's leg.

they will be capable of assembling atoms into any form whatsoever, including their own. This ability to multiply themselves by replication is important, because vast numbers of them will be needed to construct macroscopic entities such as television sets and rocket motors, atom by atom. They will work rather like the scanning tunnelling microscope — whose stylus can indeed pick up single atoms, move them about and place them in a chosen location. The final products will have no atom out of place. A favoured material of fabrication is diamond. Once the system is up and running, all human material needs will be instantly accessible by simple fabrication.

Even the faithful have qualms about the resulting paradise. How will human beings spend their time, when deprived by universal plenty of their central purpose (seen by Drexler as the competitive display of consumer durables)? What about military and evil applications of nanotechnology? Suppose the assemblers run wild and convert the whole world into intractable grey goo?

But to the unconverted (and it will be obvious by now that I am one of them) the vision is even less convincing. Even if it worked, nanotechnology would not be a

universal panacea. How, for example, would it reverse the spreading deserts or halt the population explosion? But the big question, of course, is whether it would work. Regis provides five main arguments in its favour: Feynman has endorsed the feasibility of arbitrarily small machines; scanning tunnelling microscopes have already manipulated single atoms successfully; the normal syntheses of chemistry in effect manipulate single atoms, and do it in trillionfold parallel into the bargain; bacteria function and replicate, so it must be possible; and the uncertainty principle and brownian motion do not seem to worry them.

These arguments deserve closer scrutiny than Regis gives them. How small can a machine, or more precisely a machine tool, be made? It has two essential components: a 'cutter' to act on the workpiece, and a 'frame' to act as a reference, move the cutter around the workpiece and absorb the cutting forces. In a standard engineering machine tool, such as a numerically controlled lathe or milling machine, the cutter is just that; but the analysis applies more generally.

A fundamental requirement of any machine tool is that the frame must be much bigger than the cutter. In a lathe or milling machine, the cutter has to apply its force over such a small area that the local stress deforms or shears the workpiece. The reaction to this force is of the same magnitude, but it must be absorbed by the frame with a deflection that is small compared with the desired resolution of the cutting action. The frame therefore needs a much larger cross-section than the cutter, so that this force represents a very small stress within it. The tool can then impose any pattern of deformation on the workpiece without change or damage to itself. Armed with overwhelming force, it can operate in a completely dictatorial engineering mode.

This requirement for a frame much bigger than the cutter extends to non-mechanical tools as well. A photofabricator, for example, has an optical 'cutter'. To focus as sharply as a wavelength, its lens (by the familiar laws of optics) must be many wavelengths across. To work as an atomic manipulator, a scanning tunnelling microscope might use an electrical potential as its 'cutter'. It would hold and release atoms by applying and removing a strong electric field across them. For this to work, its stylus and stage must be much larger than an atom. A stylus (say) a million atoms long will sustain at most a field one millionth of the one it imposes on the target atom. Now imagine making the microscope smaller and smaller, shrinking it down to a Drexler assembler and even

beyond. With every tenfold reduction in size, its internal field rises tenfold. It could perhaps reach the micrometre scale in working order. But by the time it approached the size of the target atom, the field within it would be disrupting its own atoms, and its operation would fail. Once a machine tool is so tiny that its frame cannot be much larger than its cutter, it can no longer function as a tool. Engineering must give way to chemistry, the science of single atoms or groups of atoms acting on each other.

Chemical constraints

And chemistry is a much humbler science than engineering. No chemist can impose an arbitrary design on his raw material, as an engineer can on his. Chemistry is a matter not of dictatorial programming but of cunning and luck and selection. It has to seek out favourable conditions and suitable reagents. It tries to set up molecular encounters that, by their own nature, their random thermodynamic shuffling towards the most accessible local energy minimum and global entropy maximum, will form the desired product. A great many substances and material forms can be made this way, but an infinitely greater number cannot. The successes of chemistry cannot be represented as an argument for nanotechnology.

Indeed, the nanotechnologists do not seem to realize the chemical obstacles in their path. To break a chunk of raw material into its component atoms needs a lot of energy — at least the latent heat of vaporization. And the single atoms when you have them cannot just be picked out and pushed around like so many marbles. When D. M. Eigler and E. K. Schweizer wrote the IBM logo in single atoms in a scanning tunnelling microscope, they used inert xenon atoms in an ultra-high vacuum at liquid helium temperature, dodging the main physical and chemical problems; and even then the atoms stayed put on their substrate only in certain positions and spacings. Single atoms of more structurally useful elements, at or near room temperature, are amazingly mobile and reactive. They will combine instantly with ambient air, water, each other, the fluid supporting the assemblers, or the assemblers themselves. And even if you manage to assemble them into some masterpiece of nanostyling, you had better be sure that it is everywhere an energy-minimum on the atomic scale. Otherwise it won't be stable. Its atoms will recrystallize and diffuse and react towards their own guess at such a minimum. You will have terrible problems with chemistry.

Ah yes, says the nanoenthusiast, but what about biochemistry? What about bacteria, those living exemplars of self-replication and construction on the micro and nano scales? If chemistry is so restricted and troublesome, how do they

manage? The answer, I think, is by extraordinary chemical specialization. Life seems to be an exercise in doing the best you can with the minimum of information. The genetic code does not specify the whole structure of an organism atom by atom. It provides a recipe that, if followed without too many blunders, produces a functioning creature capable of handing on much the same recipe. Over evolutionary time, those creatures have prospered whose recipes are the simplest and most robust for what they do.

The same constraint appears in the software of living behaviour. Termites, for example, collectively construct huge nests, in the same sort of way that Drexler imagines his assemblers building a telephone or a dishwasher. Philip Morrison analysed this process instructively in 1979. He showed that the nest results from a mass of termites all blindly following an algorithm with a few simple rules. Nowhere is there any plan or blueprint and no two nests are alike, but the structures always work as nests. It is clear that termite nests, like the termites themselves, have evolved as minimum-information structures. Termites whose nests required a more complicated algorithm failed to survive. Could termites be modified by genetic engineering to assemble, not crude variable nests but (to use Morrison's example) identical working astronomical telescopes? No: the enormous algorithm required would simply overwhelm them.

So it is wrong-headed to argue that, because a creature as small as a bacterium can reproduce itself, therefore a nanotechnological assembler can also replicate itself and, furthermore, will have enough spare capacity to replicate anything else you care to imagine. The chemistry of life centres on some highly specialized organic chemicals: DNA to carry information, a set of amino acids from which to form the proteins that double both as catalysts and as structural material, sugars and fats as energy-stores, and so on.

It is reasonable to guess that this particular biochemistry has survived and stabilized because it operates effectively with the absolute minimum of information. It permits and constitutes the ultimate in elegant programming. To claim that nanotechnological assemblers could work the same chemical trick with chemically unobtrusive materials such as aluminium and diamond, and replicate themselves with enough spare capacity to replicate anything else as well, is equivalent to asserting that life itself could be made to work in these materials, and with the added handicap of a vast overhead of arbitrary information.

This, of course, is chemically incredible. Nanotechnology will not be able to avail itself of the sparse, elegant, specialized programming evolved by biochemistry. It will have to work by engineering. Its

assemblers will somehow have to pack not only scanning tunnelling microscope technology or its equivalent into their bacterial dimensions, but the terabytes of brute-force coding and processing power needed to specify themselves and their products atom by atom. And I don't believe it will fit.

Regis's final argument concerns the two imprecisions that lurk in the nano-domain, the uncertainty principle and brownian motion. The uncertainty principle, set up and knocked down several times in *Nano!*, is rightly discounted. It doesn't stop atoms having a location or molecules having a structure; it merely makes their edges a bit fuzzy. Only at the subatomic level would it start to give serious trouble.

Brownian motion, however, cannot be dismissed so lightly. It raises a fundamental question that the nanoenthusiasts seem to have ignored: that of entropy. Formally, this damns their entire project. Above absolute zero, entropic considerations prevent even a single crystal, energy-minimum that it is, from ever being perfect. By the same token, the purest product of nanotechnology could never exist at room temperature "with no atom out of place", as Drexler dreams. But let us not be so unforgiving. Crystals are impressively regular structures despite their dislocations, and cells have learnt to live with imperfections in their DNA. A sound energy-minimized nanotechnological design should be able to tolerate a certain proportion of misplaced atoms without disaster. What constraints does thermodynamics place on its assembly?

Demon assemblers

The very first nanotechnologist was Clerk Maxwell's celebrated Demon, proposed in 1871. This was the little fellow who sat in the connecting pipe between two reservoirs of gas, watching the molecules pass between them. By blocking all molecules travelling one way, but allowing free passage to those going the other way, the Demon could build up a pressure difference across the two reservoirs. By letting fast molecules go one way and slow ones the other, he could build up a temperature difference. Either of these could be used to generate mechanical power, so the Demon could extract energy from a thermal system at one temperature — in contravention of the second law of thermodynamics.

Maxwell's Demon was a puzzle to physicists for many years, but has been well exorcized by now. The first blow was struck by Leo Szilard in 1929, when he pointed out that the Demon needed to obtain information about the molecules. He argued that the entropy generated by this interrogation exactly cancelled the reduction of entropy that could be achieved by exploiting the answer. This

elegant analysis initiated our modern identification of information with negative entropy; it has since been taken further and made part of computing theory. It is now clear that Maxwell's Demon cannot possibly work.

Nanotechnological assemblers look suspiciously like Maxwell's Demons. Drexler's argument that they are "programmed, like infecting a bacterium with a virus", evades important questions about their thermodynamics and information flow. How do the assemblers get their information about which atom is where, in order to recognize and seize it? How do they know where they themselves are, so as to navigate from the supply dump to the correct position in which to place it? How will they get their power, for commination to single atoms, navigation and, above all, for massive internal computing? How will they dispose of the entropy of their operations, and how much will they have to dispose of? The best modern computers still dissipate about 10^{12} times as much entropy as is theoretically needed by their information flow, so even allowing for improvements it could be quite a lot. Until these questions are properly formulated and answered, nanotechnology need not be taken seriously. It will remain just another exhibit in the freak-show that is the boundless-optimism school of technical forecasting. □

David E. H. Jones is in the Department of Chemistry, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, UK.

Proving the rule

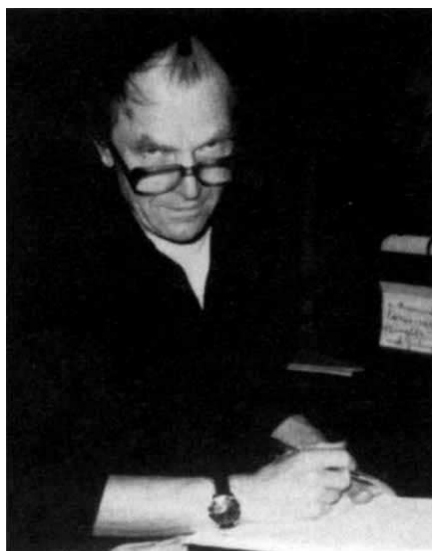
Peter Lipton

Killing Time: The Autobiography of Paul Feyerabend. By Paul Feyerabend. University of Chicago Press: 1995. Pp. 181. £18.25, \$22.95.

PAUL Feyerabend, who died in February last year, was the gadfly of the philosophy of science. Born and raised in Vienna, he fought for Germany during the Second World War, suffering bullet wounds to the spine that left him without effective use of his legs for the rest of his life. After the war, he returned to Vienna to study physics, with extensive philosophy and operatic singing on the side. Although much of Feyerabend's doctoral work was on a technical problem in classical electrodynamics, he ended up writing a philosophy thesis, for which he received a doctorate in 1951. He was then awarded a British Council scholarship to study with Wittgenstein in Cambridge, but had to find a different supervisor when Wittgenstein died the same year. He chose Karl Popper and became for a short time a card-carrying falsificationist. Soon, how-

ever, he rejected Popper's philosophy of science with a vengeance, taking it as a model of how philosophy ought not to be done. In the long run, few contemporary philosophers of science managed to avoid Feyerabend's scorn.

Feyerabend's writings, best represented by *Against Method* (New Left Books, 1975), is a verbal blitzkrieg: aggressive, fast-moving and powerful. It is also enter-



Feyerabend: science's worst enemy?

taining, infuriating and sometimes offensive. His primary target is the traditional philosophical project of explaining and defending scientific method. His strategy is to argue that any rules of scientific practice must be either so weak as to exclude nothing or persistently violated by the best scientific work. As he characteristically puts it: "The only principle that does not inhibit progress is: *anything goes*".

Feyerabend's attack on the very possibility of a non-trivial account of scientific practice comes in two stages. First he claims that the most basic methodological rules proposed by various philosophers of science were in fact broken in historical episodes that those philosophers themselves regard as models of science at its best, episodes that brought about the Copernican revolution, kinetic theory and quantum mechanics. The rules allegedly broken include those requiring that a new theory should be consistent with its predecessors or at least with the available evidence, that it should avoid *ad hoc* hypotheses and even that it should not be self-contradictory. The poor philosophers are themselves caught in a contradiction between the rules they invent and the science they admire.

The second stage of Feyerabend's attack on method deploys a variety of arguments to show that the violation of standard rules is not only a historical fact but is also essential to scientific progress. The arguments depend on two radical claims: that data are never independent of

theory and that competing theories may be logically incommensurable. Feyerabend uses these claims to argue that a scientific theory cannot be evaluated by testing it against experiment and observation alone, but only by confronting it with radical alternatives. This case for theoretical pluralism gives Feyerabend the excuse to take his attack on the philosophers into overdrive, applying it to science itself. In this scheme, effective criticism involves confronting scientific theories with all sorts of nonscientific world views, however bizarre. And there is no presumption that conventional science ought to win out in the end.

Feyerabend's defence of voodoo is unlikely to impress many readers of *Nature*, but his case against the philosophical project of defining rules of scientific practice has considerable force. He certainly scores palpable hits against particular methodological prescriptions, especially those of a Popperian stripe. This does not, however, prove the negative. Showing the failures of one or another account of scientific method is not the same as showing that no successful account is possible. Moreover, even if scientific research does not lend itself to a rule-based description, the possibility of a general account remains because rules are not the only source of generality.

On this last point, Thomas Kuhn's *Structure of Scientific Revolutions* (University of Chicago Press, 1962) provides an excellent foil to *Against Method*. Kuhn and Feyerabend hold many radical views in common: scientific theories may be incommensurable, there are no facts independent of theory, and scientific research is not governed by rules. But the perceived absence of rules does not lead Kuhn to abandon the idea of a general account of scientific research. Instead, he replaces rule-based accounts of science with one based on the power of concrete problem-solutions or exemplars to guide research in the absence of explicit rules. Feyerabend's epistemological anarchy is not the only alternative to the view that there could be a rule-book of good scientific practice.

Feyerabend's autobiography includes some philosophy, but it is mostly an indiscreet reminiscence of the life and times of a wild philosopher. It makes entertaining reading and provides plenty of material about Feyerabend's early life for anyone who wants to attempt to give a psychological explanation for his later intellectual excesses. Readers interested in the provocative arguments themselves, however, are better off thickening their skins and reading *Against Method*. □

Peter Lipton is in the Department of History and Philosophy of Science, University of Cambridge, Free School Lane, Cambridge CB2 3RH, UK.

Farewell to unreason

Paul R. Gross

The Trouble With Science. By Robin Dunbar. *Faber*: 1995. Pp. 213. £14.99, \$22.95.

ROBIN Dunbar, a psychologist and anthropologist at the University of Liverpool, England, bids us to remember not only that science is one among many products of the cognitive tool-set called 'reason', but also that "the human mind was not designed as a rational scientific mind". He provides evidence — the hammer among reason's tools — for both claims in this readable book. Not that the mind's *unreason* is despised: more often than not over the millennia it has actually been preferred. It has been advanced repeatedly, for instance, by literary scholars and social theorists, from Erasmus in the sixteenth century to modern (and postmodern) voices as various as those of Karl Marx, D. H. Lawrence, Adolf Hitler, Michel Foucault and, most recently, Václav Havel. Lawrence's "belief in the blood, the flesh, as being wiser than the intellect" has the same ancient source as Hitler's maundering to Hermann Rauschning that "*Es gibt keine Wahrheit!*".

Dunbar's "trouble with science" turns out to be one of several. He does not suggest, however, that science is anything other than the best device for getting at reality. On the contrary: he considers that the trouble-makers are detractors of science who seek to dominate opinion about science or nature, to control science policy or to deny altogether the possibility of truth. But this is no surprise: rationality, especially scientific rationality, has only recently provided any selective advantage over other modes of thinking. Evolution designed primate thought primarily for effective socialization, not for its ability to understand reality.

It has always been stylish to deny that empirical science has any particular distinction as a way of gaining knowledge about the world. And there are now certain benefits in rejecting the very possibility of distinguishing different kinds of knowledge (except, perhaps, for oneself). Irrationalism — including the trendy varieties espoused by post-positivists, some historians of science, the newer sociologists of knowledge and the prides of academic lions and lionesses doing well on identity politics — lies behind this rejection. Unfortunately, the denial is not immured in academic institutions: it titillates a public whose admiration for science when it seems useful or entertaining turns easily to dismissal or hatred when it is difficult or when myths are challenged.

Dunbar worries about the decline of

science education in the United Kingdom, where avoidance of science by the brightest students has reached scandalous proportions. In the United States, the situation is equally serious. Of course, some of the brightest students do study science; it is, for example, a requirement for admission to medical schools. But the scientific literacy of the rest — the majority — of students is of no real concern to the university. Worse still, they are taught by staff in other disciplines who are increasingly hostile to science. The scientific artlessness of graduates does harm: it has nothing to do with the funding of the Superconducting Super Collider or with the management of technology; it has everything to do with judging arguments — all arguments — and making informed decisions, whether about teaching and learning, patently false claims (such as 'alternative healing') or environmental threats, real and imagined.

Dunbar has written a strong but accessible defence of science. He has avoided technical detail and the safe self-indulgence of endnotes. (There is however an adequate and carefully selected bibliography.) Yet his points of evidence are not mere assertions; they support the merit of science as a way of finding out about the world and reveal the triviality of its fashionable dismissals. There is, in fact, a modern science — one that pre-dates the Enlightenment science invented by eighteenth-century Englishmen. The roots of this universal achievement are narrow but lie deep in human evolution. Science works by trying to find explanations, and it has always eventually succeeded — as its

record shows. Strong inference (as John Platt named it) is relatively new, whereas scientific inference, Dunbar argues, in general is not. Nor is it European, or white, or male or hegemonic. It is probably not limited to *Homo sapiens*. Negotiation creates the consensus (always temporary) of science on any question. But contrary to the belief of the socio-anthropologists of laboratory life, this negotiation is not about bandits apportioning booty. Rather, it is about what kind of evidence allows a definitive rejection or the temporary acceptance of explanations about nature.

Survival of our species, perhaps of all species, depends on our doing the best possible science and on the public's understanding of it. We seem unfortunately to be embarked on a reduction in that understanding. Dunbar's book, among others recently published or being written, is, I hope, a signal of a reaction to the systemic anti-science that has taken root, not only among the Old Right but also among the New Left. It deserves to be widely read — not least by journalists and the new academic critics of science — and to be made even more accessible by appearing in paperback as soon as possible (without the amusing misspellings). As regards science as a "way of knowing", Wittgenstein seems for once to have been clear as well as right: about that of which one cannot speak, one should shut up. □

Paul R. Gross is at the Centre for Advanced Studies, University Of Virginia, 444 Cabell Hall, Charlottesville, Virginia 22903, USA.

The New York Review version

Walter Gratzer

Hidden Histories of Science. Edited by Robert B. Silvers. *New York Review of Books*: 1995. Pp. 192. \$19.95.

If you want to have it from the horse's mouth, attend to the words of Clerk Maxwell: one used, he said, to teach the corpuscular theory of light. Now one taught the wave theory; and that was because those who believed in the corpuscular theory had died. Scientists cleave to the common currency of their discipline as to a favourite pair of socks, discarded only with reluctance when they no longer keep out the draught: this is an inalienable feature of the scientific process. Nor can the pursuit of science be separated from tenacious, even passionate, commitment to ideas. From time to time historians of science discover anew that a certain intolerance to gross heterodoxy ensues and that it sometimes impedes the advance of knowledge. But does it not far more often repel folly and keep credulity at bay?

The relationship between entrenched orthodoxy and apostasy forms the broad theme of a collection of essays solicited from his stable of science reviewers by the editor of *The New York Review of Books*. And on scientists as the enemies of science Oliver Sacks generates the strongest, though not always the most informed, opinions. He recycles some of his earlier stories — good ones, to be sure — and draws on the fates of some scientific paranoiacs to point to a moral. Georg Cantor became "floridly psychotic" on account of his persecution by a mandarin of German mathematics, Felix Klein; Boltzmann was driven to suicide by the attacks on his *confrères*, and so on. But Boltzmann, a depressive certainly, was a full professor at 25, was summoned for an audience by the Emperor Franz Josef, attracted the adulation of the younger physicists and was generally held to have vanquished his intellectual opponents, such as Ostwald, by the time his confidence failed. The career of Chandrasekhar and the rise of

relativistic astrophysics were no doubt held back by the opposition of his former patron, Eddington, who nevertheless became the most articulate apologist for Einstein and a leading exponent of the theory of the expanding Universe. Far more remarkable than Sacks's examples of resistance to assaults on the established order is, so it seems to me, the speed with which the revolution in physics in the first decades of this century took hold and the eagerness with which so many of the best physicists seized on the implications of relativity, quantum theory and uncertainty.

I also feel Sacks's thesis that Goethe's theory of colour perception, spurned by his scientific contemporaries, has been vindicated by Edwin Land's "relativity" hypothesis is stretching it a little. Everything, of course, has been prefigured at some time, just as all the prophecies of Nostradamus have been fulfilled. Salvador Dali pointed to a helix in the corner of one of his early canvases as proof that it was he who had discovered the structure of DNA. And was the phlogiston theory really, as Sacks asserts, "preposterous"? For a long time it actually answered very well, and could one not argue that it rather accurately anticipated the classification of exothermic and endothermic reactions? But now Sacks, having softened up the reader with these prods to the ribs, fells him with a kidney-punch: Freud's case histories, he declares, now confidently on home territory, are "serious science, and they embody and point to a science of the individual which is every bit as 'hard' as physics or molecular biology". (Oh, Medawar, thou shouldst be living at this hour!) Were the prevailing orthodoxies in the clinical psychiatry of only a few years ago then, such as frontal lobotomies and electroconvulsive therapy, equally firmly rooted in 'hard' science?

Jonathan Miller also ruminates, though less assertively, on anticipated discoveries — ideas that arrived too early for the existing corpus of knowledge to assimilate. Miller makes the journey from Mesmer and the myth of animal magnetism that caused such a *frisson* in eighteenth-century Europe, until it was debunked by the famous commission of *savants* set up by Louis XIV ("*L'imagination fait tout, le magnétisme nul*"), to the English neurologist John Hughlings Jackson and the function of the cerebral cortex. He takes a tortuous, though interesting, route to his final conclusion — that the mind is aware of only a small part of itself, for it contributes unbidden to much of what we perceive, and that this same message fell on deaf ears for a century and more.

Richard Lewontin's essay is a rebuttal of determinism and an assertion of the influence of the environment on development and form. (Unspoken in the background are the many exasperated words that Lewontin has already expended on the great red herring of genes and intelligence.) If the environment alters the organism, he says, the organism also impinges on its environment. Therefore the desire that man should leave the

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

This parody of evolutionary progress appeared in London's *Times* last year during the week of a UK train strike and the announcement in *Nature* of the discovery of *Australopithecus ramidus*, hailed as the earliest-known hominid.

world in the condition in which he found it is misplaced. Lewontin's views will have little appeal for the environmentalists, for he implicitly discounts the unique powers that man has developed for inflicting such calamitous abuse on his surroundings. Man in our time, Julian Huxley believed, had become the cancer of the Earth; Lewontin will have none of that.

Stephen Jay Gould is on good form in his homily on how imagery or, as he calls it, iconography, subverts thought. The evolutionary tree or the very form of the conventional 'cone' of biological diversity, drawing the eye upwards, has helped to perpetuate the identification of evolution (a term that Darwin felt was inherently misleading) with improvement. This view of evolution as a purposeful process, by which the organism improves its performance, was expounded by such influential popularizers as C. H. Waddington in the 1930s and 1940s, and is still evidently lodged in the public perception. Gould shows how merely opening out the cone diagram for the life-forms of the Burgess Shale creates a different and, he persuades us, truer impression. Perhaps the common representation derives from the traditional portrayal of human pedigrees (from *pied de grue*, or a crane's foot, with toes deviating slightly from the vertical).

The most absorbing of the pieces in this collection is (to me) that of Daniel Kevles. He calls his story "Pursuing the unpopular: a history of courage, viruses

and cancer", and his heroes are a succession of intrepid researchers — Peyton Rous, John Bittner and Ludwik Gross prominent among them, and in our own time the pioneers of RNA viruses, reverse transcriptase and oncogenes — who breached the prevailing orthodoxy and in time supplanted it, in the face of much scepticism and sometimes hostility. Gross was perhaps the most remarkable in his faith and persistence, for he conducted his experiments on the transfer of leukaemia between mice in the time that he could spare from his work as a hospital doctor over seven long years with no grant, until, not without a slice of luck, he prevailed. As for Bittner, convinced as he was that the agent of breast cancer in his mice was a virus, he found it prudent to refer to it as a "factor", because the cancer virus hypothesis had been discredited, and denounced indeed, in a report by the US Surgeon General. Had Bittner allowed the word to slip out in a publication he would, he believed, have been branded a crank and forfeited all prospects of grant support.

Kevles suggests that, in today's climate, Gross, whose success in transferring the leukaemia virus turned out to be a function of the strain of mice that he happened to be using and was for some time not reproduced by other laboratories, might have been hauled before the Holy Office (of Research Integrity, that is) and cast into darkness. Of course, such devotion to a belief as Gross's is rare in science, as in other human endeavours. Consider, for instance, Trollope's evaluation of seekers after public office: "If membership of the Treasury bench were confined to men who believe the world to be triangular, then there would arise from Ministerial aspirants a great assertion of triangularity." But Kevles does not draw the facile conclusion vociferated by so many critics of science. Here is his moral: "What permitted the pioneers eventually to prevail was to a significant extent their professional courage, imagination, and persistence. Yet it was also the tolerance and pluralism of the basic biomedical research system — the tolerance of deviant ideas and the pluralism that provides niches (large like Rous's and Temin's or small like Gross's) in which the ideas have a chance to flourish."

Robert Silvers, the editor, has inspired a rewarding collection of good writing. Even where the substance inflames, the literary excellence assuages. Let's hear it for *The New York Review of Books*. □

Walter Gratzer is at the MRC Muscle and Cell Motility Unit, King's College, 26-29 Drury Lane, London WC2B 5RL, UK.

Reductive megalomania

John L. Casti

Nature's Imagination: The Frontiers of Scientific Vision. Edited by John Cornwell. Oxford University Press: 1995. Pp. 212. £16.99, \$23.

WHILE watching Hollywood go through its annual exercise in narcissism recently, the circus more commonly known as the Academy Awards, I couldn't help thinking how much more entertaining a cinematic rendition of this volume would be than watching a bunch of rednecks torment a simpleton like Forrest Gump. All the components of a long-running blockbuster are here in abundance: an eternal storyline filled with tension and conflict; heroes and heroines valiantly arguing completely contradictory points of view, a mild-mannered director who keeps the story moving along but never intrudes, an almost mediaeval setting amidst the cloistered halls of academe and even a few good one-liners to leaven the sombre tone of much of the dialogue. How could it fail? Happily, it can't. This book is the outgrowth of a meeting held in Cambridge, England, in September 1992, at which 15 scholars from widely varying disciplines took up the question of the adequacy of scientific reductionism as a general strategy for attacking questions about the natural and human worlds. The end result is a collection of intellectual vignettes welded together into a drama that anyone interested in complex systems and the philosophy of science will want to read and savour.

The leitmotif of the stories told here is the age-old problem of reductionism: can we always understand a given system by decomposing it into simpler parts? A large part of the theatre in this book revolves about the fact that the question seems to be understood quite differently depending on who is telling the tale. Physicists, by and large, tend to see the question as asking whether all biological and mental events are ultimately reducible to the properties of matter and energy. Just about everyone else (this reviewer included) sees the problem of reductionism in more algorithmic terms, revolving about whether a given system can always be 'reduced' to basic 'atoms' whose properties can then be used to re-assemble the behaviour of the original system.

One of the good guys in the white hats in this drama is the mathematical physicist Roger Penrose, who argues eloquently for an anti-reductionistic view of both mathematics and mathematical physics. Addressing his well-known view that the brain — or at least the mind — can never be duplicated by a simple rule-following machine,

Penrose employs everything from the global topological properties of impossible figures to the nonlocality of quantum theory to buttress this claim. Further ammunition, at least for belief that mathematics is not reducible to a set of deductive rules, is provided by Gregory J. Chaitin, who demonstrates the existence of perfectly sensible mathematical propositions whose truth or falsity we will never be able to prove. As Chaitin has termed the situation, "sometimes things are true [or false] for just no reason whatsoever".

The ultimate reductionistic fantasy is undoubtedly the so-called 'theory of everything' (TOE) that physicists have for centuries been claiming is just around the corner. The cosmologist John Barrow out-



lines the reasons why such a theory may be necessary — but far from sufficient — for a description of the Universe and its contents, emphasizing the important point that we never observe a law of nature but only its consequences. Moreover, any law of nature, including a TOE, is simply a compression of what we can see. But Chaitin's work shows that it is impossible to know if any compression is really the best possible. So it would never be possible to know if any TOE really was what it claimed to be. Barrow also provides a bit of much-needed, although probably unplanned, comic relief when he states that "the greatest discovery of twentieth-century science is that the Universe is expanding". I wonder if there is anyone other than a cosmologist who could make such a remarkable claim with a straight face.

The leader of the gang of hardcore, dyed-in-the-wool reductionists in the black hats is the chemist P. W. Atkins, who, in a ten-page outburst that is simultaneously a paean to reductionistic science and a no-holds-barred frontal assault on all religions, tries to convince us that science is capable of answering every question about the world, the flesh and the devil that could ever arise in any enquiry. His method? Reductionism, what else? In the second most hilarious line of this saga, Atkins

describes his position thus: "Science has never encountered a barrier that it has not surmounted". He terms this the "omni-competence of science".

One of the volume's twin heroines is the philosopher Mary Midgley, who quite correctly captures the tone of this Atkinsesque-style of thinking in the title of her presentation, "Reductive Megalomania". Levelling her guns at the reductive approach as a general attitude, Midgley blasts the view that reduction is value-neutral; rather, she claims, it is always part of some positive propaganda campaign. Generally speaking, this propaganda is directed at showing how something (usually physics) is more 'fundamental' than something else (usually everything). She counters this brand of megalomania by asking the reader to try translating a factual sentence such as "George was allowed home from prison at last on Sunday" into the deeper truths of physics that the Atkinses of the world believe underlie it. Midgley's assertion is that for such a translation, all the social concepts such as 'Sunday', 'home', 'allowed' and 'prison' would have to vanish and be replaced by terms involving the interactions of groups of elementary particles (or, if you're a TOE man or woman, strings).

The other heroine of the piece is the philosopher and computer scientist Margaret Boden. She takes as her theme the brand of reductionism that sees humans as being 'reduced' to mere machines, first by the steam engines of the Industrial Revolution and more recently by the possible emergence of intelligent computing machines. In a passionately and cogently argued assessment of this view, Boden concludes that the spectre of artificial intelligence does not reduce our respect for human minds. Quite the contrary in fact: it increases our respect for human qualities and helps us to understand how the mind is at all possible.

A number of supporting players — Gerald Edelman, Hao Wang, Oliver Sacks, W. F. Clocksin and Paul and Patricia Churchland — make their appearance in this drama, somewhere between the alpha of a holistic Universe and the omega of strong reductionism à la Atkins. In one way or another, these players contribute not only to the all-star character of the cast but also to the overwhelming impression that reductionism as a sole-source supplier of scientific methodology is now all but buried in the avalanche of evidence pouring in from aardvarkology to zymurgology. As the wise old sage Freeman Dyson states in the introductory chapter, "Science is an art form and not a philosophical method". This book is one of the strongest testaments to that belief. Read and enjoy. □

John L. Casti is at the Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA.

Symbiosis evolving

David L. Hawksworth

Evolution by Association: A History of Symbiosis. By Jan Sapp. Oxford University Press: 1994. Pp. 255. \$49.95, £37.50 (hbk); \$24.95, £18.95 (pbk).

SYMBIOSIS as a concept has evolved multifariously since its inception. The term now embraces associations at vastly different scales, from those of organelles within eukaryotic cells to the interplay of organisms in Gaia. Current usage also encompasses associations that are parasitic as well as mutualistic, concepts in themselves open to differing interpretations.

To appreciate how this situation originated necessitates an historical perspective. Jan Sapp has striven to provide such a back-cloth through extensive bibliographic research, supplemented by discussions with individuals, the list of which reads as one of potential nominees for a putative World Prize in Symbiology. But he has not shied away from presenting controversial and personal interpretations.

Despite the significance of symbiosis in evolution, the topic has received short shrift in discussions of the development of evolutionary biology. Indicative of this is its omission from Ernst Mayr's *The Growth of Biological Thought* (1982).

Contrary to most credits, the term originated in an 1877 paper by Albert B. Frank (1839–1900) on lichen anatomy. Frank used the word in a neutral way irrespective of the roles of the organisms living together, paralleling its use by H. Anton de Bary (1831–88) in an address the following year. De Bary did not refer to Frank in the printed version of that presentation, but he cited Frank's paper elsewhere. That two scientists addressing similar issues, and once based in universities 35 km apart (Halle and Leipzig) — a point not noted by Sapp — should arrive independently at the same term would be extraordinary. Such enthusiastic researchers must have discussed such issues. We may never know who first coined the word, but Frank had a particular interest in terminology, proposing "mycorrhiza" in 1885.

The strategies of 'commensalism', 'mutualism' and 'parasitism', already distinguished by the Belgian zoologist Pierre-Joseph van Beneden in 1873, were rapidly subsumed under the general term 'symbiosis'. Van Beneden opposed natural selection as the major evolutionary force, and before the turn of the century symbiosis was increasingly treated as synonymous with mutualism and a concept in opposition to the Darwinian 'struggle for existence'. Symbiosis assumed a sociopolitical dimension, moved out of the scientific comfort-zone, and its study fell into disrepute. Symptomatically, the status of the topic conspired with that of women in sci-

ence to divert Beatrix Potter's energies from studies of lichen symbiosis.

By 1899 the mutualistic concept was being extended by Herbert Spencer to the integration of the living world as a superorganism. This in part foreshadowed the debate started by James Lovelock's *Gaia* (1979), addressed in Sapp's penultimate chapter. Again, scientific hypotheses proposed for testing, and which in this case highlighted the importance of multidisciplinary approaches, moved into a sociopolitical and pseudoscientific arena — and thus tarnished its scientific standing.

Symbiogenesis, the role of symbiosis in the evolution of the eukaryotic cell, was explored independently by the Russian scientists K. S. Merechovsky and A. S. Famintsyn, and later also by B. M. Kozopolianski, in the first decades of this century. Contrary to the claims of some modern symbiologists, Sapp shows this to have been a sideline rather than a major element of Soviet evolutionary thought. Paul

Portier's *Les Symbiotes* (1918), published in Paris and suggesting that all organisms other than bacteria were formed of associations between different kinds of creatures, also became embroiled in debates on microorganisms as disease agents, the resulting controversy further reducing the subject's standing. It was not until advances in ultrastructural and subsequently molecular techniques from the 1970s that the 'serial endosymbiosis theory' of the eukaryotic cell regained scientific respectability. In the second edition of *Symbiosis in Cell Evolution* (1993), which appeared too late for Sapp to cite, Lynn Margulis considered that only spirochaete origins for undulipodia remained the truly "questionable aspect".

In keeping with issues still under debate, Sapp highlights the treatment of symbiosis in *The Science of Life* (H. G. Wells, J. S. Huxley and G. P. Wells; 1929–30) — although referencing the Doubleday (New York) edition rather than the Amalgamated Press (London) original with an identical text but different chapter headings and pagination. These authors stressed that although superficially many associations between organisms might appear to be mutually beneficial, more detailed examination often made it difficult "to distinguish service from slavery".

The "unscientific aura of teleology" was



CRYING game — the Asian moth *Hypochrosis baenzigeri* makes an elephant cry and then drinks its tears. From *Elephants* by Clive Spinae, an authoritative and comprehensive account of the origins, evolution, anatomy, ecology, behaviour and conservation of this long-time associate of humans. Nicely illustrated with both line drawings and photographs, the book concludes by looking at the domestication of the elephant, past and present. T&AD Poyser, £25.

considered by the editors of the Society for General Microbiology's symposium volume *Symbiotic Associations* (1963) to "disguise a fundamentally parasitic relationship". Held at the Royal Institution in London, the meeting included presentations of elegant experimental work on the nature of certain symbiotic associations and stimulated critical scientific work that developed exponentially over the next two decades, mainly through the energy of David Smith's group in Oxford. I would have welcomed more analysis of the results of such investigations, of experimental work on isolated partners, and the insights from synthesis experiments and ultrastructural studies.

For lichens, new evidence of 'controlled parasitism' at the cellular level has prompted the question whether mutualism should be defined at the individual cell

level or in terms of increased fitness of the partners at the social level. Within lichen associations in which neither partner occurs in a free-living form, the symbiosis cannot be other than mutually beneficial at the social level. As lichens are such a key element in the development of the original concept of 'symbiosis', I would have enjoyed commentary on this topical scaling controversy.

Overall, the book is a delightful historical complement to the available texts describing symbiotic associations and a scholarly overview of the conceptual aspects of the subject. It will be enjoyed both by symbiologists and by those interested in the history of biological thought. □

David L. Hawksworth is at the International Mycological Institute, Bakeham Lane, Egham, Surrey TW10 9TY, UK.

Juicy beetle and other stories

S. J. Simpson

Bugs in the System: Insects and their Impact on Human Affairs. By May R. Berenbaum. Addison-Wesley: 1995. Pp. 337. \$25 (hbk); £19.95 (pbk).

WE entomologists are a frustrated lot. It is blindingly obvious that insects are both beautiful and amazing, so why doesn't everybody else realize this? Granted, the aesthetic appeal of many insects is perhaps a tad difficult to convey to those poor souls without an entomocentric view, but insects *are* amazing. And how do we entomologists try to convert our long-suffering colleagues, friends and, in fact, anyone else unfortunate enough to enquire politely what we do for a living? Why, we tell them amazing things about insects. May Berenbaum has provided us

with the ultimate source of incredible entomological facts. As I read the book I folded back the corner of each page containing an extraordinary piece of entomologia. My copy will not close now. There is enough here to keep entomological evangelists in stories for generations to come.

Did you know that the production of 1 kilogram of honey involves 10 million visits to individual flowers and bees flying the equivalent of 10 times around the Earth? Or that the extraordinary abundance of chamber pots excavated from the Aztec capital was due to human urine being used as a mordant for cochineal dye? Or that the larvae of some flea species attach to the anus of the adult and imbibe their faeces? Or that there is a club in Washington, DC, that specializes

in organically produced mealworm wontons and cricket peanut-butter cups? Or that an elephant-dung pat in East Africa was removed in 2 hours by 16,000 dung beetles? Do you know what cantharidiasis or autothysis are? Or that...?

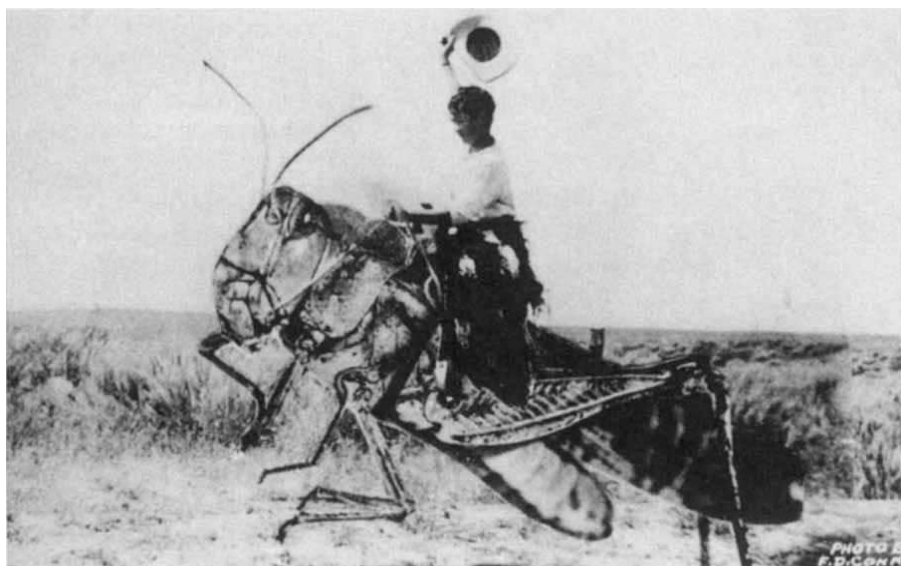
Berenbaum has provided a comprehensive account of the ways in which insects have influenced human affairs, along with salient features of insect behaviour, physiology, social organization and life history. Aspects as diverse as the importance of insects in art, agriculture, food, fly fishing, forensic medicine, mythology, warfare, social structure and even human evolution are covered. Implicating insects in the evolution of the opposable thumb (for social grooming) may be over-entomocentric, but emphasizing their impact on other aspects of human society and biology clearly is not.

For example, the Second World War was the first ever war in which louse-transmitted typhus did not kill more people than the number who died in battle. In 1812 Napoleon embarked on his attempt to conquer Russia with an army of 600,000 men. A year later only 3,000 remained, most of the rest having died of typhus. Thanks to its demographic effects, the Black Death contributed greatly to the decline of the manorial system in Europe. The automobile was encouraged as a non-polluting alternative to horse-drawn vehicles, which were associated with dung-breeding, disease-transmitting fly populations.

This fascinating compendium is written in an appropriately light-hearted style. There are plenty of bad puns (when referring to the extreme competition among carrion-feeding insects, for example, the author says: "It's a dead-dog-eat-dead-dog world out there") and, inevitably, in the interests of a balanced coverage, there is a lot of sex and defecation, so the book will appeal to all biologists. Whether, with its mass of detail, it will hold the attention of lay readers, particularly those with a delicate constitution, is perhaps doubtful. But then again, they will receive the information second-hand from us entomologists on aeroplanes and buses and at dinner parties all over the globe — whether or not they want to hear it.

Finally, I must just mention the case in 1970 when a hotel owner was sued by a guest "who, while attempting to get rid of a cockroach which crawled up her thighs, got her legs entangled in a bedspread and stumbled and fell over". Cantharidiasis? That is the infestation of the rectum by dung-feeding beetles. And autothysis? Explosive defecation in which the abdomen is blown off. □

S. J. Simpson is at the University Museum, University of Oxford, Parks Road, Oxford OX1 3PW, UK.



Leap of imagination — a novelty postcard from the turn of the century.

Forgetting the facts of life

Stuart Sutherland

Rewriting the Soul: Multiple Personality and the Sciences of Memory. By Ian Hacking. Princeton University Press: 1995. Pp. 336. \$24.95, £19.95.

REGARDLESS of their familiarity with the Gospel according to St Matthew, psychiatrists should adopt the motto "Seek and ye shall find". Until 1972, there were only about a hundred recorded cases of multiple personality — indeed, the classification was not listed in the official US psychiatric diagnostic manual until 1980. But from the mid-1970s, American psychiatrists began using this diagnosis more and more frequently, until now the United States contains tens of thousands of 'multiple personalities', with hospitals devoted entirely to their treatment and a learned society dedicated exclusively to their study. In the rest of the world — including the United Kingdom — there are still only a handful of cases. How can this be? Do American psychiatrists suffer from a form of mass hysteria, has there been a sudden change in the symptomatology of their patients, have they in the past simply failed to diagnose the disorder — if it is a disorder — or are they merely finding what they are looking for?

All four elements probably played a part in the surge of 'multiples', as they are now known in the jargon of the trade, but there are several further causes. One is the recent fascination with child abuse, particularly sexual abuse. For reasons Ian Hacking does not adequately explain, many psychiatrists believe that a multiple personality is the result of child abuse. If all multiples have been abused, making this diagnosis increases the supposed prevalence of child abuse. Moreover, once an illness becomes headline news, more and more people believe they are suffering from it. Forty years ago myalgic encephalomyelitis (ME, better known as 'yuppie flu') did not exist nor did attention-deficit disorder in children: now they are common. Both the psychiatrist and his patients are influenced by diagnostic fads. The psychiatrist can present original case studies of new disorders, while the patient has an excuse for misbehaving, which may and indeed has absolved her (most multiples are women) from a sentence for murder committed when a different personality was in control.

Finally, every profession tries to justify itself by claiming arcane knowledge

opposed to common sense: Freud rose to fame by constructing a brilliantly implausible theory, psychiatrists exhibit their skills by spotting and (according to Hacking) creating multiples, while social workers are adept at uncovering nonexistent child abuse. Part of the popularity of the last concept comes from people's morbid delight in persecuting imagined evil-doers — be they child abusers, witches or (in Nazi Germany) Jews. Once the evil is publicized, there is no shortage of people claiming to have experienced it, as witness the crop of patients who alleged they had been the victims of Satanic rituals: the prevalence of such rituals can be gauged by the fact that at one time just as many patients claimed to have been abducted by aliens.



Double Portrait by Hans von Marées, 1863.

Playing the part of a multiple ensures attention, but are multiples merely playing a part? To the extent to which switches between different alters (alternative personalities) are involuntary, they are not: moreover, the amnesia under one alter (usually a subordinate one) about what happened when under another one appears to be genuine, although it is only autobiographical facts that are forgotten not how to talk or dress.

The history of multiple personality and its connection with child abuse is a good story, but Hacking does not tell it well. His prose is convoluted and he often takes several pages to make an obvious point. For example, he distinguishes psychiatrists' conscious knowledge from their methods of practice, of which they may not be consciously aware. But he confounds the reader by using the term "knowledge" when what is relevant is psychiatrists' beliefs, and he has a lengthy digression on Foucault who, although not the first to make the distinction, succeeded in making it more obscurely than any of his predecessors. Having made the point, Hacking makes nothing of it. He raises many hares but catches few. It is

unclear whether he believes that the sole cause of multiple personality is child abuse: he presents no evidence either way. And although it is clear that many recovered memories are fantasies (often implanted by psychotherapists eager to keep up with fashion), he equivocates on the prevalence of child abuse. He does, however, hint that it may be harmless, at least if no cruelty is involved: as a child, Louis XIII of France repeatedly had his genitals stroked in public. In the past, children in the West were treated much more savagely than those of today. They were beaten, sent up chimneys, starved and disease ridden. Although cruelty to children cannot be condoned, it does not necessarily lead to multiple personalities or any other forms of mental disorder.

Despite its subtitle, there is nothing in *Rewriting the Soul* about the science of memory, which in any case throws no light whatsoever on multiple personality. Hacking believes that only if we define a person through the continuity of his or her memory could the concept of multiple personality have arisen and he contrasts this view of personality with thinking of the 'soul'. But the concept of different souls existing in the same body is just as easy to grasp as different sets of memories, and indeed it goes back a long time — that is what exorcism was about. On the other hand, Hacking is occasionally sage: he argues that the publicity given to child abuse will almost certainly increase its practice, just as suicides portrayed on television are followed by an increased number of real suicides (unfortunately he fails to quote the definitive reference on this point).

In psychiatry, both treatment and diagnosis are often based on caprice rather than reason. Hacking describes Pierre Janet's method of helping women who had suffered a childhood trauma, by hypnotizing them into believing that the traumatic event never occurred. He also describes the modern technique of dealing with post-traumatic stress disorder — persuade the patient to relive the event. But he does not call attention to the fact that these therapeutic methods are exact contraries, although he seems to accept the dubious cures claimed by both sets of clinicians.

Perhaps the most shocking aspect of the whole story is that not a single scientific study of the aetiology or treatment of multiple personality is reported: presumably none has taken place. Clinical judgements are notoriously unreliable and despite all the seeking and finding, we know no more about this condition than did Morton Prince a hundred years ago. The human mind remains a more elusive entity than the most distant galaxy. □

Stuart Sutherland is in the Laboratory of Experimental Psychology, University of Sussex, Brighton, BN1 9QG, UK.

The mind of an unhappy genius

A. Rupert Hall

Descartes: An Intellectual Biography. By Stephen Gaukroger. Oxford University Press: 1995. Pp. 499. £25, \$35.

WHAT is an 'intellectual biography'? It must surely be more than either a life-history or an analysis of achievements set out in chronological order, or even a discussion of those features of a person's work that seem to stem directly from his or her experience. To understand Wordsworth's lines "It is a beauteous evening, calm and free/The holy time is quiet as a Nun/Breathless with adoration", we need not only to know of hidden events in the poet's life but also to penetrate some way into his character and such social matters as attitudes to love-children. So intellectual biography is perhaps a richer form when emotional content is high and certainly it requires a range of private material to make the connection between a man or a woman and the work.

We can never know seventeenth-century figures such as Harvey, Descartes or Newton as well as we know Wordsworth. The human materials for biographical depth do not exist. For Descartes (1596–1650), details of the first 30 years of his life are few; thereafter, several volumes of correspondence still leave obscurities in his development and personal relationships with others. We know that he fathered a child of a servant, we suspect that he suffered (especially as a young man) more than one psychological crisis and we may infer (from episodes of unreasonable intellectual jealousy, for example) that he was unstable and insecure: his life was that of an unhappy wanderer. But really we know Descartes only from his writings, including the correspondence that was so intimately related to his mathematical and philosophical works. Perhaps that should be enough.

A peculiar issue in the case of Descartes is the way in which the condemnation of Galileo in 1633 broke his life in two. No satisfactory explanation of his tenderness on this point is known. Most of what Descartes had composed before that event was published only posthumously, if at all, although much was rewritten into later works printed between 1637 and

1649. Descartes was thus 41 — and of course already well known in intellectual circles — before a word of his was printed, and perhaps the most important part of his intellectual biography lies in the story of unfinished drafts and projects, and in the way in which these contributed to his brief productive period and to the posthumous books. As the author emphasizes, however, Descartes' new intellectual explorations during the later period of his life were in "legitimatory metaphysics". In natural philosophy his creative work was over. "The condemnation raised the question of how natural philosophy was to be legitimated, and in particular how it could be raised above the level of the hypothetical." Galileo's misfortune led directly to philosophy three most famous words: *Cogito ergo sum*.

The connection is traced — and a major theme of this book epitomized — as follows: "The problem that Descartes started off with was that of building up a natural philosophy, mechanism, as a basis for a microcorpuscularian, hydrostatically/hydrodynamically modelled physics whose fortunes were inextricably tied to those of Copernicanism."

Stephen Gaukroger is an experienced and erudite writer on Descartes. He has a complete mastery of the recent literature, and is at home equally in metaphysics, epistemology and natural philosophy. He offers reasoned argument against many conventional criticisms of Descartes, for example that his physics was derived from metaphysics. (His defence should, however, be read in the light of a later claim that Descartes "is able to establish (metaphysically) the unique legitimacy of a particular way of pursuing natural philosophy without raising a single natural-philosophical question".)

But as the quoted passages may indicate, this is not an easy book and there are many stones to catch the unguarded toe. What, for instance, is "hyperbolic doubt"? Historians of mathematics may shudder a little at a quasi-dismissal of the *Géométrie* as "cobbled-up" from Descartes' earlier writings on mathematics; yet his use of the "mesolabe compass" is well treated. Some passages must be pretty opaque to readers imperfectly familiar with the Cartesian canon and earlier scholarly discussions. Of more structural importance is the absence of a systematic assessment of Descartes' character as a man and a philosopher — out of place in a philosophical study of Descartes' writings but appropriate in a

biography. True, there are accounts (generally unflattering) of his treatment of, for example, Isaac Beeckman or Fermat. Descartes, totally unreceptive to criticism, was assured of the truth of an intellectual system that, as Gaukroger explains, was tightly interlocked (if not completely consistent on every point), asserting once that if the velocity of light were not infinite, as he supposed, then all his philosophy must be false. He often abused those who could not see as clearly as himself.

Yet for us too that philosophy is often impenetrable. Descartes was on the far side of the divide created by Newton between philosophy and science. He was also unashamedly theological in his metaphysics and epistemology — to the point, more than once, of having his cake and eating it (as he did also in making Earth a planet while not, by his definition, allowing it to move). Yet Gaukroger is surely right to treat Descartes as the first unambiguously modern philosopher. He steers firmly, if not always simply, through the complexities created by the pressures to which Descartes' thinking was subject; ending with Descartes' skull, chopped off 16 years after his death by a Swedish officer and now in the Musée de l'Homme in Paris. □

A. Rupert Hall is at 14 Ball Lane, Tackley, Oxfordshire OX5 3AG, UK.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

New Journals

This year, *Nature's* annual new journals review supplement will appear in the issue of 21 September. Publishers and learned societies are invited to submit journals for review, taking note of the following criteria:

- Journals that first appeared during or after June 1993 and issued at least four separate numbers by the end of April 1995 will be considered.
- Journals covering any aspect of science are eligible, although those dealing with clinical medicine and pure mathematics are excluded, as are publications of abstracts.
- Frequency of publication must be at least three times a year. The main language used must be English. Translation journals in English are, of course, eligible.
- Deadline for submission is the end of May.

When submitting journals for review, please send at least four different issues (the first, the most recent and any two others) of each title, together with full details of subscription rates (personal and institutional) and frequency of publication to: Peter Tallack, *Nature*, Macmillan Magazines Ltd, Porters South, Crinan Street, London N1 9SQ, UK. Tel: +44 (0)171 843 4567.